TUGAS AKHIR – KS184822

# KLASIFIKASI KADAR TOKSISITAS DAN PROTEKSI RADIASI MENGGUNAKAN *NEURAL NETWORK ENSEMBLE* DAN *RANDOM FOREST* PADA SENYAWA TERPENTING GEN SUPRESOR P53

DIAN RIZKY MAULINA
NRP 062115 4000 0091

Dosen Pembimbing
Dr. rer. pol. Heri Kuswanto, S.Si., M.Si.

PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA. KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019

TUGAS AKHIR – KS184822

# KLASIFIKASI KADAR TOKSISITAS DAN PROTEKSI RADIASI MENGGUNAKAN *NEURAL NETWORK ENSEMBLE* DAN *RANDOM FOREST* PADA SENYAWA TERPENTING GEN SUPRESOR P53

DIAN RIZKY MAULINA
NRP 062115 4000 0091

Dosen Pembimbing
Dr. rer. pol. Heri Kuswanto, S.Si., M.Si

PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019

FINAL PROJECT – KS184822

# CLASSIFICATION FOR CITOTOXICITY AND RADIOPROTECTIVE USING NEURAL NETWORK *ENSEMBLE* AND *RANDOM FOREST* FOR THE IMPORTANT COMPOUND OF SUPRESOR GENE P53

DIAN RIZKY MAULINA
NRP 062115 4000 0091

Supervisor
Dr.rer.pol. Heri Kuswanto, S.Si., M.Si.

UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019

*(Halaman ini sengaja dikosongkan)*

## LEMBAR PENGESAHAN

**KLASIFIKASI KADAR TOKSISITAS DAN PROTEKSI RADIASI MENGGUNAKAN *NEURAL NETWORK ENSEMBLE* DAN *RANDOM FOREST* PADA SENYAWA TERPENTING GEN SUPRESOR P53**

**TUGAS AKHIR**

Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Statistika
pada
Program Studi Sarjana Departemen Statistika
Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember

Oleh :
**Dian Rizky Maulina**
NRP. 062115 4000 0091

Disetujui oleh Pembimbing:
**Dr. rer. pol. Heri Kuswanto, S.Si., M.Si.** ( )
NIP. 19720925 199803 2 001

Mengetahui,
Kepala Departemen Statistika

**Dr. Suhartono**
NIP. 19710929 199512 1 001

SURABAYA, JULI 2019

*(Halaman ini sengaja dikosongkan)*

# KLASIFIKASI KADAR TOKSISITAS DAN PROTEKSI RADIASI MENGGUNAKAN *NEURAL NETWORK ENSEMBLE* DAN *RANDOM FOREST* PADA SENYAWA TERPENTING GEN SUPRESOR P53


**Nama Mahasiswa : Dian Rizky Maulina**
**NRP : 062115 4000 0091**
**Departemen : Statistika-FMKSD-ITS**
**Dosen Pembimbing : Dr. rer. pol. Heri Kuswanto, S.Si., M.Si.**


## Abstrak


*Kanker merupakan penyakit akibat pertumbuhan tidak normal akibat sel-sel jaringan tubuh yang berubah menjadi sel kanker. Salah satu pengobatan pada kanker adalah radioterapi, namun memiliki efek samping yaitu membunuh sel-sel normal disekitar sel kanker. Penggunaan radioterapi secara terus-terus dengan radiasi yang tinggi juga akan berpengaruh kepada pasien kanker. Sehingga, diperlukan adanya penjustifikasian klasifikasi senyawa-senyawa penting gen supresor p53 dengan akurasi yang tinggi dengan menggunakan senyawa-senyawa yang diperoleh dari penelitian sebelumnya yang memperoleh perankingan 20 besar senyawa penting gen supresor p53 yang baik digunakan untuk radioterapi. Klasifikasi kadar toksisitas dan radio proteksif dengan menggunakan metode neural network ensemble lebih baik dibandingkan dengan menggunakan random forest karena didapatkan nilai akurasi yang lebih tinggi.*



***Kata kunci: Gen Supresor p53, High Dimensional Data, Kanker, Neural network Ensemble, Proteksi Radiasi, Random forest, Toksisitas***

*(Halaman ini sengaja dikosongkan)*

***CLASSIFICATION OF CITOTOXICITY AND RADIOPROTECTIVE USING NEURAL NETWORK ENSEMBLE AND RANDOM FOREST IN IMPORTANT COMPOUNDS OF SUPRESSOR P53 GEN***


**Name : Dian Rizky Maulina**<br>
**Student Number : 062115 4000 0091**<br>
**Department : Statistics**<br>
**Supervisor : Dr. rer. pol. Heri Kuswanto, S.Si., M.Si.**



**Abstract**

*Cancer is a disease caused by abnormal growth due to the cells of the body's tissues that turn into cancer cells. One treatment for cancer is radiotherapy, but it has a side effect of killing normal cells around the cancer cells. Continuous by using radiotherapy with high radiation will also affect to the cancer's patients. So it takes a drug compound that is able to control it. Thus, in this study the classification of important compounds of the p53 suppressor gene was carried out with high accuracy using compounds obtained from previous studies which obtained a ranking of the top 20 important compounds of the p53 suppressor gene which is well used for radiotherapy. Classification of citotoxicity and radioprotective by using neural network ensemble is more better than using random forest because it brings higher accuracy values than than random forest have.*



***Keywords : Cancer, High Dimensional Data, Neural network Ensembles, Radiation Protection, Random forest, Supressor genes p53, Toxicity***

*(Halaman ini sengaja dikosongkan)*

# KATA PENGANTAR

Puji syukur penulis panjatkan atas rahmat dan hidayah yang diberikan Allah SWT sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul “Klasifikasi Kadar Toksisitas Dan Protesi Radiasi Menggunakan *Neural Network Ensemble* Dan *Random Forest* Pada Senyawa-Senyawa Terpenting Gen Supresor P53” dengan lancar.

Penulis menyadari bahwa Tugas Akhir ini dapat terselesaikan tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Kedua orang tua, atas segala do’a yang diberikan kepada penulis demi kesuksesan dan kebahagiaan penulis.
2. Dr. Suhartono selaku Ketua Departemen Statistika dan Dr. Santi Wulan Purnami, M.Si. selaku Ketua Program Studi Sarjana yang telah memberikan fasilitas, sarana, dan prasarana.
3. Dr. Bambang Widjanarko Otok, S.Si, M,Si. selaku dosen wali selama masa studi yang telah banyak memberikan saran dan arahan dalam proses belajar di Departemen Statistika.
4. Dr. rer. pol. Heri Kuswanto, S.Si., M.Si. selaku dosen pembimbing yang telah meluangkan waktu dan dengan sangat sabar memberikan bimbingan, saran, dukungan serta motivasi selama penyusunan Tugas Akhir.
5. Dr. Suhartono dan Dr. Santi Putri Rahayu, S.Si., M.Si selaku dosen penguji yang selalu sabar dalam mengomentari serta memberikan masukan dan saran dalam penyelesaian Tugas Akhir.
6. Sahabat-sahabat penulis, Shelma Hajar, Laras Raditia, Andina Primadini, Naomi Puspita, Dwita Ajeng Martidhania, Anggi Wahyu, Adita Krisdiyantoro, Harira Yunita, Rovita Bharadela, dan Crisda Yana yang selama ini telah membantu, mendukung, menyarankan, dan

mendengarkan keluh kesah penulis selama mengikuti berbagai kegiatan pada masa perkuliahan berlangsung.

7. Teman-teman seperjuangan TA, Rosikhu Ilmi, Charles Rudiyanto, Ika Nur Laili, dan Dwinda Intan yang selama ini telah berjuang bersama dan saling memberikan semangat
8. Teman-teman tempat banyak cerita tersalurkan, Shindi Shella, Desintya Rachma, Dissa Rahmayani, Arlandio, Habib Jazuli, dan Geng Kopi (Aprilia, Natasya Dea, Novita, Zikra, Anindya, Bina) yang selalu membersamai penulis dalam suka dan duka.
9. Kakak-kakak angkatan, mba Erlin, mas Reynaldi, mba Neni Alya yang sangat membantu dalam proses penyusunan Tugas Akhir ini, mulai dari awal hingga terselesaikannya Tugas Akhir ini.
10. Seluruh dosen Statistika ITS yang telah memberikan ilmu dan pengetahuan yang tak ternilai harganya, serta segenap karyawaan Departemen Statistika ITS.
11. Teman-teman Statistika ITS ∑26 angkatan 2015, yang selalu memberikan dukungan kepada penulis selama ini.
12. Semua teman, relasi dan berbagai pihak yang tidak bisa penulis sebutkan namanya satu persatu yang telah membantu dalam penulisan laporan ini.

Besar harapan penulis untuk mendapatkan kritik dan saran yang membangun sehingga Tugas Akhir ini dapat memberikan manfaat bagi semua pihak yang terkait.

Surabaya, Juli 2019

Penulis

# DAFTAR ISI

Halaman

**LEMBAR PENGESAHAN**.......**Error! Bookmark not defined.**
**ABSTRAK** ........................................................................ v
**ABSTRACT** .....................................................................vii
**KATA PENGANTAR**...........................................................ix
**DAFTAR ISI**.......................................................................xi
**DAFTAR GAMBAR**..........................................................xiv
**DAFTAR TABEL**...............................................................xvi
**DAFTAR LAMPIRAN**........................................................xix
**DAFTAR NOTASI**............................................................xxiv
**BAB I PENDAHULUAN**...................................................... 1
1.1 Latar Belakang............................................................ 1
1.2 Rumusan Masalah ....................................................... 6
1.3 Tujuan Penelitian.......................................................... 6
1.4 Manfaat Penelitian........................................................ 6
1.5 Batasan Masalah........................................................... 7
**BAB II TINJAUAN PUSTAKA**............................................ 9
**2.**1 Statistika Deskriptif...................................................... 9
2.2 *Neural network* ............................................................ 9
2.3 *Random forest*............................................................ 13
2.4 Evaluasi Ketepatan Klasifikasi.................................. 14
2.5 Kanker ....................................................................... 16
2.6 Toksisitas................................................................... 17
2.7 Radioterapi ................................................................ 18
**BAB III METODOLOGI PENELITIAN**............................ 21
3.1 Sumber Data .............................................................. 21
3.2 Variabel Penelitian .................................................... 21
3.3 Langkah Analisis........................................................ 24
**BAB IV ANALISIS DAN PEMBAHASAN** ........................ 29
4.1 Karakteristik Data *Citotoxicity* dan *Radioprotective*.. 29

4.2 Klasifikasi dengan *ANN Ensemble*............................31
4.2.1 Ilustrasi Metode *ANN Ensemble*....................31
4.2.2 Klasifikasi Senyawa Penyusun Obat Kanker Menggunakan ANN *Ensemble*..........38
4.3 Klasifikasi Menggunakan *Random forest*...................57
4.3.1 Ilustrasi *Random Forest*................................57
**BAB V KESIMPULAN DAN SARAN**................................63
5.1 Kesimpulan.................................................63
5.2 Saran.......................................................64
**DAFTAR PUSTAKA**.............................................65
**LAMPIRAN**...................................................69

*(Halaman ini sengaja dikosongkan)*

# DAFTAR GAMBAR

Halaman

**Gambar 2.1** Ilustrasi ANN dan Jaringan Syaraf Manusia ......10
**Gambar 2.2** Arsiterktur *Neural network* MLP........................13
**Gambar 2.3** Metastasis dan Apoptosis Gen p53.....................17
**Gambar 3.1** Penentuan *threshold* Pada Kelas *Cytotoxicity* dan *Radioprotective* ..........................................24
**Gambar 3.2** Diagram Alir Penelitian.....................................26
**Gambar 4.1** Proses Gen terhadap Mutagen ............................29
**Gambar 4.2** Gen Supresor p53 ..............................................30
**Gambar 4.3** Proporsi Jumlah Kelas Respon Data *Citotoxicity* dan *Radioprotective* ...................................30
**Gambar 4.6** llustrasi *Artificial Neural network Ensemble*......37
**Gambar 4.7** llustrasi Proses Pada *Input Layer*........................38

*(Halaman ini sengaja dikosongkan)*

# DAFTAR TABEL

Halaman


**Tabel 2.1** Tabulasi Silang Ketepatan Klasifikasi ....................15
**Tabel 2.2** Kategori Kebaikan Model ......................................16
**Tabel 3.1** Struktur Data Penelitian untuk Toksisitas ($Y_1$) .......21
**Tabel 3.2** Struktur Data Penelitian untuk Proteksi Radiasi ($Y_2$) ..............................................................21
**Tabel 3.3** Variabel Penelitian ................................................22
**Tabel 4.1** Probabilitas *Ensemble* Pertama data *citotoxicity* dengan Fungsi Aktifasi *Softmax* Pada *Subspace* 12 ...........................................................34
**Tabel 4.2** Hasil Klasifikasi ANN pada *Citotoxicity* dengan 11 *Ensemble* Untuk 12 *Subspace* Menggunakan Aktifasi *Softmax* ................................................35
**Tabel 4.3** Tabulasi Silang Data Ilustrasi.................................37
**Tabel 4.4** Arsitektur ANN *Ensemble* .....................................39
**Tabel 4.5** *Probability* Setiap *Subspace* untuk Fungsi Aktivasi *Softmax*.................................................41
**Tabel 4.6** Tabulasi Silang pada *ANN Ensemble* Fungsi Aktivasi *Softmax*.................................................43
**Tabel 4.7** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Softmax Threshold* Optimum..............................44
**Tabel 4.8** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Softmax Threshold* 0.5.......................................44
**Tabel 4.9** *Probability* Setiap *Subspace* untuk Fungsi Aktivasi *Sigmoid* ................................................45
**Tabel 4.10** Tabulasi Silang pada Model ANN *Ensemble* Fungsi *Sigmoid*...................................................47
**Tabel 4.11** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Sigmoid Threshold 0.3*.......................................48

**Tabel 4.12** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Sigmoid Threshold 0.5* ....................................... 48
**Tabel 4.13** *Probability* Setiap *Subspace* untuk Fungsi Aktivasi *Softmax* ............................................... 49
**Tabel 4.14** Tabulasi Silang pada Model ANN *Ensemble* Fungsi *Softmax* .................................................. 51
**Tabel 4.15** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Softmax Threshold 0.625* ................................... 52
**Tabel 4.16** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Softmax Threshold 0.5* ....................................... 52
**Tabel 4.17** *Probability* Setiap *Subspace* untuk Fungsi Aktivasi *Sigmoid* ............................................... 53
**Tabel 4.18** Tabulasi Silang pada Model ANN *Ensemble* Fungsi *Sigmoid* .................................................. 55
**Tabel 4.19** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Sigmoid Threshold 0.625* ................................... 56
**Tabel 4.20** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Sigmoid Threshold 0.5* ....................................... 56
**Tabel 4.21** *Probability* dan Kelas Prediksi Data *Citotoxicity* . 58
**Tabel 4.22** *Probability* dan Kelas Prediksi ............................ 60
**Tabel 4.23** Tabulasi Silang pada *Random forest* .................... 61
**Tabel 4.24** Ukuran Kebaikan Model *Random forest* ............. 61

*(Halaman ini sengaja dikosongkan)*

# DAFTAR LAMPIRAN

Halaman

**Lampiran 1.** Data Citotoxity .................................................. 69
**Lampiran 2.** Data Radioprotective........................................ 70
**Lampiran 3.** Syntax Random Forest ..................................... 71
**Lampiran 4.** Syntax ANN *Ensemble* .................................... 72
**Lampiran 4.** Syntax ANN *Ensemble* (lanjutan)..................... 73
**Lampiran 5.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 12 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax* ........................... 77
**Lampiran 6.** Alokasi Variabel Prediktor Pada *AN Ensemble* 13 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax* ........................... 78
**Lampiran 7.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 14 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax* ........................... 79
**Lampiran 8.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 15 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax* ........................... 80
**Lampiran 9.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 16 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax* ........................... 81
**Lampiran 10.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 17 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax* ........................... 82
**Lampiran 11.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 18 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax* ........................... 83
**Lampiran 12.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 19 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax* ........................... 84

**Lampiran 13.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 12 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*............................85
**Lampiran 14.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 13 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*............................86
**Lampiran 15.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 14 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*............................87
**Lampiran 16.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 15 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*............................88
**Lampiran 17.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 16 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*............................89
**Lampiran 18.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 17 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*............................90
**Lampiran 19.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 18 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*............................91
**Lampiran 20.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 19 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*............................92
**Lampiran 21.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 12 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax*........93
**Lampiran 22.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 13 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax*........94
**Lampiran 23.** Alokasi Variabel Prediktor Pada *ANN*

*Ensemble* 14 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax* ....... 95
**Lampiran 24.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 15 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax* ....... 96
**Lampiran 25.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 16 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax* ....... 97
**Lampiran 26.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 17 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax* ....... 98
**Lampiran 27.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 18 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax* ....... 99
**Lampiran 28.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 19 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax* ..... 100
**Lampiran 29.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 12 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid* ..... 101
**Lampiran 30.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 13 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid* ..... 102
**Lampiran 31.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 14 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid* ..... 103
**Lampiran 32.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 15 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid* ..... 104
**Lampiran 33.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 16 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid* ..... 105

**Lampiran 34.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 17 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid*......106
**Lampiran 35.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 18 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid*......107
**Lampiran 36.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 19 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid*......108

# DAFTAR NOTASI

***Neural Network***

| | |
|---|---|
| $x_{i(k)}$ | Variabel input sebanyak *p,* (i=1,2,…,*p*) |
| $\hat{y}_{(k)}$ | Nilai dugaan dari variabel *output* |
| | Indeks pasangan data *input-target* |
| $w_{ji}^{h}$ | Bobot dari *input* ke-*i* yang menuju *neuron* ke-*j* pada lapis tersembunyi |
| $b_{j}^{h}$ | Bias pada *neuron* ke-*j* pada lapis tersembunyi |
| $f_{j}^{h}$ | Fungsi aktifasi di *neuron* ke-*j* pada lapis tersembunyi |
| $w_{j}^{0}$ | Bobot dari *neuron* ke-*j* di lapis tersembunyi yang menuju *neuron* pada lapis *output* |
| $b^{0}$ | Bias pada *neuron* di lapis *output* |
| $f^{0}$ | Fungsi aktifasi pada *neuron* di lapis *output* |

*(Halaman ini sengaja dikosongkan)*

# BAB I
# PENDAHULUAN

## 1.1 Latar Belakang

Kanker merupakan penyakit tidak menular yang saat ini mendapatkan perhatian yang serius di dunia. Kanker menjadi masalah kesehatan yang serius karena menjadi penyebab kematian sebanyak 8,2 juta pada tahun 2012 (RI, 2013). *World Health Organization* memperkiraan kematian akibat penyakit kanker akan terus bertambah dari 8,2 juta kematian pada tahun 2012 diproyeksikan menjadi 11,4 juta pada tahun 2030. Kanker sendiri adalah penyakit yang disebabkan oleh pertumbuhan sel-sel jaringan tubuh yang tidak normal. Sel-sel kanker akan berkembang dengan cepat dan tidak terkendali serta akan terus membelah diri. Kemudian, sel kanker akan menyusup ke jaringan sekitarnya (invasif) dan akan terus menyebar melalui jaringan ikat, darah, serta menyerang organ-organ penting dan syaraf tulang belakang. Dalam keadaan normal, sel tubuh hanya akan membelah diri jika ada penggantian sel-sel yang telah mati dan rusak. Sebaliknya sel kanker akan terus membelah walaupun tubuh tidak memerlukannya. Hal ini mengakibatkan terjadinya penumpukan sel baru yang disebut tumor ganas. Penumpukan sel baru tersebut akan mendesak dan merusak jaringan normal sehingga mengganggu organ yang ditempatinya (TimCancerHelps, 2010).

Pada tahun 2018, *International Agency for Research on Cancer* (IARC) mendata bahwa kanker menyebabkan 9,6 juta kematian pada tahun 2018 dan menjadi penyebab kematian nomor dua di dunia setelah kardiovaskuler. Berdasarkan data *World Health Organization* kasus kanker secara global telah meningkat menjadi 18,1 juta kasus baru dan 9,6 juta kematian di tahun 2018. Kematian terbanyak akibat penyakit kanker terjadi di Asia dengan jumlah persentase sebesar 57,3% yang lebih tinggi daripada Eropa dengan persentase sebesar 20,3%. Sedangkan di Indonesia, pada tahun 2018, berdasarkan *International Agency for Research on Cancer* (IARC) memiliki jumlah kasus kanker baru sebanyak

348.809 kasus dengan total kematian akibat penyakit kanker sebesar 207.210 jiwa. Kasus-kasus baru kanker di Indonesia secara keseluruhan didominasi oleh jenis kanker payudara, kanker servix, kanker paru-paru, kanker usus, dan kanker hati.

Beberapa terobosan inovasi untuk mengobati kanker telah banyak diciptakan, salah satunya yaitu radioterapi dengan menggunakan kadar dosis radiasi tinggi untuk membunuh sel kanker dan mengecilkan ukuran tumor. Penggunaan dosis rendah pada radioterapi biasanya digunakan pada *X-Ray* untuk melihat kondisi dalam tubuh. Sedangkan pada dosis tinggi, radioterapi akan membunuh sel kanker atau melambatkan pertumbuhannya dengan merusak DNA sel tersebut. Radiasi tidak hanya membunuh atau menghambat pertumbuhan sel kanker, tetapi juga dapat berimbas pada kerusakan sel-sel normal disekitarnya (NationalCancerInstitute, 2017). Untuk mengatasi efek radioterapi tersebut, Ariyasu *et al.* (2014) mendesain proteksi radiasi dengan mencari komponen senyawa yang berhubungan dengan protein gen p53. Fungsi dari gen p53 adalah mencegah replikasi sel pada sel yang rusak secara genetik melalui penghentian siklus sel pada fase G1 atau interfase, sehingga sel mempunyai waktu untuk melakukan perbaikan. Selain itu, gen p53 berfungsi untuk mencetuskan apoptosis apabila kerusakan sel cukup luas dan terjadi kegagalan pada proses perbaikan. Apabila terjadi mutasi pada gen p53, maka akan mengakibatkan disregulasi gen p53 sehingga terjadi kegagalan apoptosis dan sel yang rusak akan terus mengalami replikasi, hingga menyebabkan terjadinya suatu keganasan.

Dari penelitian yang dilakukan oleh Ariyasu *et al* pada tahun 2014 diperoleh 84 senyawa yang diduga bagus untuk dijadikan *radioprotector* yang kemudian dari masing-masing senyawa diberi dua percobaan yaitu pada sel normal dan sel yang terkena radiasi gamma. Pada percobaan pertama dilakukan pemberian senyawa pada sel normal, percobaan ini bertujuan untuk mengukur toksisitas. Percobaan kedua dilakukan pemberian senyawa pada sel yang telah terkena radiasi sinar gamma (10 Gy), percobaan ini

bertujuan mengukur fungsi proteksi radiasi. Indikator yang digunakan pada kedua percobaan tersebut adalah dengan melihat tingkat kematian sel. Senyawa komponen yang diperlukan untuk *radioprotector* yang baik adalah suatu komponen dari senyawa yang memiliki tingkat kematian sel yang rendah pada kadar toksisitas dan tingkat kematian sel yang tinggi pada fungsi radioproteksi. Penelitian oleh Matsumoto *et al.* (2016) yang melanjutkan penelitian Ariyasu *et al.* (2014) yaitu tentang prediksi proteksi radiasi dan toksisitas menggunakan *Random forest* (RF) dan *Support Vector Machine* (SVM) dan didapatkan hasil bahwa *Random forest* lebih baik digunakan dalam memprediksi toksisitas dengan nilai AUC 77.8% sedangkan SVM baik digunakan dalam memprediksi proteksi radiasi dengan nilai AUC 64.4%. Penelitian lain dilakukan oleh Kimura *et al*. (2017) dengan membandingkan hasil ketepatan klasifikasi senyawa untuk optimasi proteksi radiasi dan toksisitas dengan menggunakan *random forest* (RF), *support vector machine* (SVM), *extreme gradient boosting* (XGB), dan K-*nearest neighbor* (kNN). Kemudian, penelitian yang dilakukan oleh Matohashi *et al.* (2018) yang melanjutkan penelitian Ariyasu *et al.* (2014) memperoleh hasil perankingan 20 senyawa gen supresor p53 yang baik digunakan untuk radioterapi dengan menggunakan *Random forest* dan *Support Vector Regression*. *Random forest* pada peneltian tersebut digunakan sebagai *feature selection*, sedangkan *support vector regression* digunakan sebagai pemodelan senyawa. Sehingga perankingan senyawa-senyawa penting gen supresor p53 didapatkan dari hasil prediksi kadar toksisitas dan proteksi radiasi menggunakan *support vector regression.* Pada penelitian Matohashi *et al*. (2018), terlebih dahulu membuat model regresi dari kadar toksisitas dan proteksi radiasi dengan menggunakan *support vector regression* dan membuat *experimental ranking* dimana perankingan tersebut digunakan untuk membandingan hasil penelitian dengan opini ahli klinik medis.

Data yang digunakan dalam penelitian Matohashi *et al.* (2018) merupakan *high dimensional data* dimana jumlah variabel

yang diamati sangat besar dan jumlah pengamatan yang dilakukan jauh lebih kecil dari jumlah variabel, sehingga dilakukan *feature selection* dalam menangani *high dimensional data*. Matohashi *et al.* (2018) dalam penelitiannya menyatakan bahwa opini dari ahli kimia medis lebih baik daripada *experimental ranking* yang dibuat. Hal tersebut didasarkan pada hasil perankingan senyawa-senyawa gen supresor p53 pada *predicted ranking* yang hanya beberapa saja yang sama dengan senyawa-senyawa pada *experimental ranking* dimana hanya ada tiga senyawa pada *experimental ranking* yang berada pada sepuluh besar senyawa di *predicted ranking*. Dalam perankingan senyawa-senyawa gen supresor p53 pada penelitian Matohashi *et al.* (2018) dilakukan dengan menggunakan metode kombinasi *pareto ranking* untuk menemukan pareto optimal dari optimalisasi multiobjektif. Optimalisasi multiobjektif memiliki fungsi multipel objektif dan biasanya berhubungan dengan *trade off* dimana fungsi radioproteksi dan toksisitas berhubungan dengan *trade off* tersebut.

Penanganan *high dimensional data* terdiri dari dua cara pendekatan yaitu mengurangi dimensi pada dataset atau dengan mengaplikasikan metode yang independen terhadap dimensional data. Pada penelitian ini menggunakan *ensemble* sebagai basis dalam klasifikasi. Metode klasifikasi dengan basis ensamble yang digunakan dalam penelitian ini adalah *neural network ensemble. Neural network* digunakan oleh Devi *et al* (2016) pada pengklasifikasian kanker serviks yang didapatkan bahwa *neural network* dapat mendeteksi kanker serviks pada stadium awal sehingga dapat menurunkan angka kematian pada wanita. Selain itu, penelitian oleh Jong & Ki (2010) menggunakan *neural network ensemble* untuk memprediksi kebangkrutan dengan *ensemble* yang digunakan adalah *bagging* dan *boosting.* Dengan menggunakan *bagging* dan *boosting,* akurasi yang didapatkan pada sepuluh topologi yang dibuat meningkat serta mengurangi masalah *overfitting.* Selain itu, *bagging* dan *boosting* juga mengurangi *error test* sebesar 8,60% dan 16,97%. Sehingga *bagging* dan *boosting* dapat dijadikan alat yang efektif untuk meningkatkan performansi

*neural network* pada kebangkrutan. Pada tahun 2003, Won & Sung mempresentasikan teknik menggunakan *ensemble* dari klasifikasi *neural network* yang dipelajari dari karakteristik berkorelasi negatif untuk secara akurat mengkategorikan kanker dan memperkirakan fungsi dari teknik yang diusulkan dengan menggunakan tiga patokan dataset. Dari pengujian membuktikan bahwa klasifikasi *ensemble* dengan karakteristik berkorelasi negatif memberikan tingkat pengakuan terbaik pada dataset.

Selain menggunakan *neural network ensemble*, metode pengklasifikasian lainnya adalah *Random forest. Random forest* merupakan metode klasifikasi yang sangat populer untuk membentuk model *ensemble decision trees* dari subruang data (Breiman, 2001). Hasil penelitian menunjukkan bahwa *random forest* dapat meningkatkan nilai akurasi dalam pengklasifikasian dalam tipe *high dimensional data.* Selain itu, menurut Huang *et al*. (2012), *random forest* juga mampu menjadi model *ensemble* di mana setiap komponen *classifier* dibangun dari subruang data yang mampu memodelkan kelas di subruang. Selain dapat digunakan pada *high dimensional data*, *random forest* juga dapat digunakan untuk mengatasi *imbalance data*. Chen *et al*. (2004) mengemukakan dua cara dalam menangani *highly imbalance data* dengan menggunakan *random forest*, yaitu dengan menggunakan *weighted random forest* dan *balanced random forest*. Kedua cara tersebut mampu meningkatkan akurasi prediksi kelas minor dan lebih baik daripada algoritma yang ada.

Berdasarkan uraian diatas, penelitian ini menggunakan metode *neural network ensemble* dan *random forest* untuk mengetahui performansi klasifikasi senyawa-senyawa terpenting gen supresor p53 yang telah terbentuk pada penelitian yang telah dilakukan oleh Matohashi *et al*. (2018). Hasil klasifikasi dari *neural network ensemble* akan dibandingkan dengan hasil klasifikasi menggunakan *random forest.* Pada penelitian ini tidak melakukan *preprocessing data* dan *feature selection* karena kedua proses tersebut sudah dilakukan pada penelitian sebelumnya. Evaluasi ketepatan klasifikasi yang digunakan adalah dengan

membandingkan nilai akurasi total, spesifisitas, dan sensitifitas tertinggi dan dipilih metode terbaik. Perbandingan kedua metode tersebut diharapkan mampu memberikan hasil yang baik.

## 1.2 Rumusan Masalah

Permasalahan yang timbul dalam perankingan kadar toksisitas dan proteksi radiasi adalah klasifikasi dilakukan secara multivariat. Sehingga diperlukan evaluasi keefektifan pada senyawa-senyawa terpenting gen supresor p53 yang telah dihasilkan oleh penelitian Matohashi *et al*. (2018). Oleh karena itu, pada penelitian ini dilakukan klasifikasi secara univariat menggunakan metode *neural network ensemble* dan *random forest* pada 20 senyawa terpenting yang telah didapatkan oleh penelitian Matohashi *et al*. (2018). Sehingga didapatkan hasil akurasi yang tinggi dan perankingan senyawa yang lebih optimal.

## 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah diuraikan di atas, tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

1. Mengetahui karakteristik data *citotoxicity* dan *radioprotective*.
2. Mendapatkan evaluasi hasil ketepatan klasifikasi dari masing-masing metode untuk mendapatkan metode yang paling baik dalam mengklasifikasikan senyawa gen supresor p53.

## 1.4 Manfaat Penelitian

Berdasarkan tujuan yang ingin dicapai, adapun manfaat yang diharapkan dari hasil penelitian ini adalah sebagai berikut.

1. Bagi bidang kedokteran, diharapkan penelitian ini dapat memberikan kontribusi dalam validasi senyawa untuk obat kanker berdasarkan tingkat toksisitas dan proteksi radiasi.
2. Bagi dunia pendidikan, penelitian ini dapat menjadi referensi dalam melakukan penelitian selanjutnya mengenai

pemodelan dan klasifikasi senyawa untuk obat kanker dengan metode *Neural network Ensemble dan Random forest*.

## 1.5 Batasan Masalah

Batasan masalah pada penelitian ini adalah tidak melakukan *pre-processing* data dan *feature selection* karena kedua proses tersebut telah dilakukan pada penelitian sebelumnya.

*(Halaman ini sengaja dikosongkan)*

# BAB II
# TINJAUAN PUSTAKA

Tinjauan pustaka yang terdapat dalam penelitian ini merupakan penjelasan dari metode yang digunakan diantaranya statistika deskriptif, *artificial neural network ensemble, random forest,* dan ketepatan klasifikasi.

## 2.1 Statistika Deskriptif

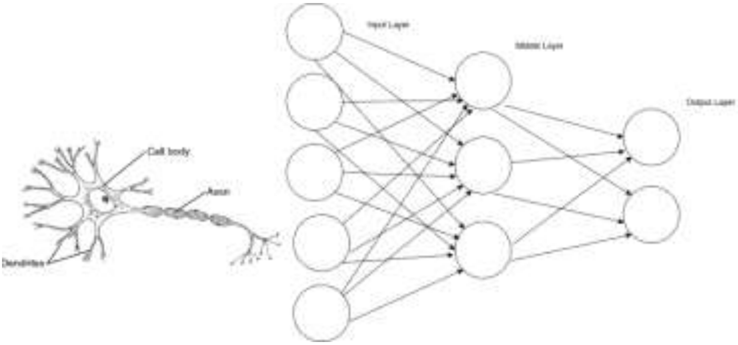
Statistika deskriptif merupakan sebuah metode statistika yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Statistika deskriptif hanya memberikan informasi mengenai data yang dipunyai dan sama sekali tidak menarik inferensia atau kesimpulan apapun. Statistika deskriptif mampu menyajikan kumpulan data yang tersedia dengan ringkas dan rapi serta dapat memberikan informasi inti dari kumpulan data yang ada.

Informasi yang dapat diperoleh dari statistika deskriptif ini antara lain ukuran pemusatan data dan ukuran penyebaran data (Walpole, 1993). Terdapat banyak macam bentuk penyajian statistika deskriptif seperti tabel, diagram batang, diagram garis, diagram *pie*, histogram, *scatter plot* dan lain sebagainya dengan harapan informasi yang disampaikan akan lebih mudah dipahami oleh pembaca. Dalam penelitian ini, statistika deskriptif yang akan digunakan adalah *mean*, maksimum dan minimum.

## 2.2 *Neural network*

*Neural network* adalah model yang terinspirasi oleh bagaimana neuron dalam otak manusia bekerja. Tiap neuron pada otak manusia saling berhubungan dan informasi mengalir dari setiap neuron tersebut (Agrawal & Agrawal, 2015). *Neural network* mempelajari pola-pola dan hubungan data yang jumlahnya sangat besar dan bersifat non-linier dengan menyaring data, mencari hubungan, membangun pola dan mengoreksi kesalahan pola itu sendiri berkali-kali. *Neural network* merupakan sistem pemroses informasi yang memiliki karakteristik mirip dengan jaringan syaraf biologi dalam memproses informasi seperti otak manusia terdiri

dari sejumlah neuron yang melakukan tugas sederhana. Karena adanya keterhubungan antar neuron, maka otak dapat melakukan fungsi pemrosesan yang cukup kompleks. Pemrosesan informasi pada manusia bersifat adaptif, yang artinya hubungan antar neuron terjadi secara dinamis dan selalu memiliki kemampuan untuk mempelajari informasi-informasi yang belum diketahui sebelumnya. Dalam bidang statistik, *neural network* mampu mengolah data dalam jumlah besar dan dapat memiliki akurasi yang tinggi dalam melakukan prediksi (Sarle, 1994). Ilustrasi *neural network* dan jaringan syaraf pada manusia terlihat pada gambar berikut.




**Gambar 2.1** Ilustrasi ANN dan Jaringan Syaraf Manusia

Tiap neuron menerima *input* dan melakukan operasi perkalian dengan sebuah pembobot, menjumlahkannya (*weighted sum*), dan menambahkan bias. Hasil dari operasi ini akan dijadikan parameter dari fungsi aktivasi yang akan dijadikan *output* dari neuron tersebut.

Bentuk arsitektur *neural network* (NN) yang secara umum paling banyak digunakan dalam aplikasi di bidang teknik atau rekayasa adalah *Multi Layer Perceptrons* (MLP). Dalam pemodelan statistik, MLP dapat dipandang sebagai suatu kelas yang fleksibel dari fungsi-fungsi nonlinear. Secara umum, model ini bekerja dengan menerima suatu vektor dari input $x$ dan

kemudian menghitung suatu respon atau output $\hat{y}(x)$ dengan memproses (*propagating*) $x$ melalui elemenelemen proses yang saling terkait. Elemen-elemen proses tersusun dalam beberapa lapis (*layer*) dan data input, $x$ , mengalir dari satu lapis ke lapis berikutnya secara berurutan. Dalam tiap-tiap lapis, input-input ditransformasi kedalam lapis secara nonlinier oleh elemen-elemen proses dan kemudian diproses maju ke lapis berikutnya. Akhirnya, nilai-nilai output $\hat{y}$, yang dapat berupa nilai-nilai skalar atau vektor, dihitung pada lapis output (Suhartono, 2007). Dalam arsitektur MLP dengan satu lapis tersembunyi yang terdiri dari $q$ unit *neuron* dan lapis *output* yang hanya terdiri dari satu unit *neuron,* nilai-nilai respon atau *output* $\hat{y}$ dihitung dengan persamaan berikut.

$$\hat{y}_{(k)} = f^0\left[\sum_{j=1}^{q}\left[w_j^0 f_j^h\left(\sum_{i=1}^{p} w_{ji}^h x_{i(k)} + b_j^h\right) + b^0\right]\right] \tag{2.1}$$

dengan,

$x_{i(k)}$ : variabel input sebanyak $p$, ($i$=1,2,...,$p$)

$\hat{y}_{(k)}$ : nilai dugaan dari variabel *output*

$k$ : indeks pasangan data *input-target* ($X_{i(k)}$,$Y_{(k)}$), $k$=1,2,...,$n$

$w_{ji}^h$ : bobot dari *input* ke-$i$ yang menuju *neuron* ke-$j$ pada lapis tersembunyi, ($j$ = 1,2,...,$q$)

$b_j^h$ : bias pada *neuron* ke-$j$ pada lapis tersembunyi ($j$= *1,2,...,*q)

$f_j^h$ : fungsi aktifasi di *neuron* ke-$j$ pada lapis tersembunyi

$w_j^0$ : bobot dari *neuron* ke-$j$ di lapis tersembunyi yang menuju *neuron* pada lapis *output*

$b^0$ : bias pada *neuron* di lapis *output*

$f^0$ : fungsi aktifasi pada *neuron* di lapis *output*

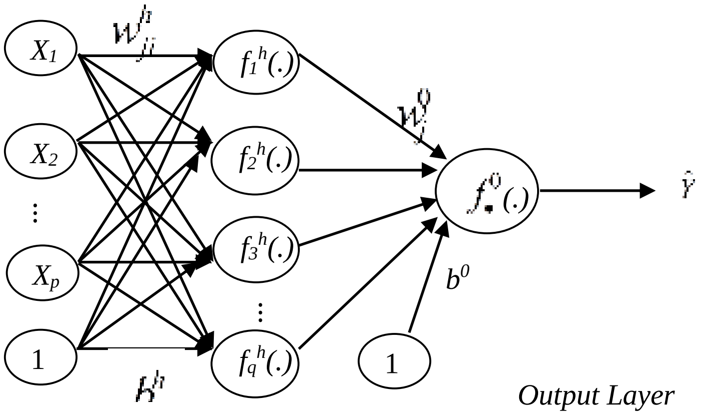
Fungsi aktivasi yang akan digunakan dalam penelitian ini adalah fungsi aktivasi *softmax* dan *sigmoid*. Persamaan matematis untuk kedua fungsi aktivasi dapat ditulis sebagai berikut.

$$f(x) = \frac{1}{1+e^{-x}} \tag{2.2}$$

$$f(x_i) = \frac{\exp(x_i)}{\sum_{k=1}^{p} \exp(x_k)} \tag{2.3}$$

i = 1,2,3, ..., p

Fungsi aktivasi *sigmoid* digunakan ketika menangani masalah *binary classification*, sedangkan *softmax* digunakan saat menangani *multiple classification* (Bishop, 2006). Pada *neural network* terdapat dua kategori utama arsitektur jaringan tergantung pada jenis hubungan antara neuron, yaitu *feed forward nerual network* dan *recurrent neural network*. Jika tidak ada *feedback* dari *output* pada neuron terhadap input seluruh jaringan, maka jaringan disebut sebagai sebuah *feed forward nerual network*. Sebaliknya, jika ada umpan balik seperti itu, yaitu sinaptik sambungan dari output terhadap masukan (input sendiri atau masukan neuron lain), maka jaringan disebut *recurrent neural network* (Bishop, 2006). Arsitektur *neural network* merupakan pola hubungan antar neuron. Neuron yang memiliki pola bobot dan fungsi aktivasi yang sama akan dikelompokkan dalam lapisan (*layer*) yang sama. Biasanya informasi yang dirambatkan dari lapisan *input* ke lapisan *output* akan melewati lapisan tersembunyi (*hidden layer*). Berikut ilustrasi arsitektur *neural network* yang akan digunakan dalam penelitian.




**Gambar 2.2** Arsiterktur *Neural network* MLP

## 2.3 *Random forest*

*Random forest* merupakan pendekatan *ensemble* yang dapat juga dianggap sebagai prediktor *nearest neighbor*. Metode *random forest* dikembangkan dari suatu metode, yaitu *classification and regression tree* (CART). Pengembangan metode CART dilakukan dengan metode *bootstrap aggregating* (*bagging*) dan *random feature selection* (Breiman, 2001). Analisis metode *random forest* dilakukan pada sekumpulan pohon yang membentuk hutan (*forest*). *Subspace* data metode ini terdiri dari n observasi dan p variabel prediktor dengan prosedur sebagai berikut :

1. Melakukan penarikan contoh acak berukuran n dengan pemulihan pada gugus data (fase *bootstrap*).
2. Dengan menggunakan contoh *bootstrap*, pohon dibangun sampai mencapai ukuran maksimum. Pembangunan pohon dilakukan dengan menerapkan random *feature selection* pada setiap proses pemilihan pemilah, yaitu *m* variabel prediktor

dipilih secara acak dimana *m* << *p*, lalu pemilah terbaik dipilih berdasarkan *m* variabel prediktor tersebut.

3. Ulangi langkah 1 dan 2 sebanyak *k* kali, sehingga terbentuk sebuah hutan yang terdiri atas k pohon.

Pada *random forest*, hasil prediksi k pohon digunakan sebagai prediksi respon suatu observasi dengan cara digabungkan. Hasil prediksi *random forest* adalah berdasarkan kategori atau kelas yang paling sering muncul sebagai hasil prediksi dari k pohon klasifikasi (*majority vote)*. Pada setiap iterasi *bootstrap*, terdapat sekitar sepertiga amatan gugus data asli yang tidak termuat dalam contoh *bootstrap*. Amatan tersebut disebut data *out of bag*. Data *out of bag* tidak digunakan untuk membangun pohon tetapi dijadikan data validasi pada pohon yang bersesuaian. *Error out of bag* bergantung pada korelasi antar pohon dan kekuatan masing masing pohon dalam *random forest* (Breiman, 2001). Ketika korelasi dinaikkan maka akan meningkatkan error *out of bag*. Sedangkan, jika kekuatan pohon dinaikkan maka akan menurunkan error *out of bag*. Saat melakukan *random forest,* ukuran contoh peubah penjelas (*m*) yang digunakan sangat mempengaruhi korelasi dan kekuatan tersebut. Meningkatkan nilai m akan meningkatkan keduanya, begitu juga sebaliknya. Penggunaan *m* yang tepat akan menghasilkan *random forest* dengan korelasi antar pohon cukup kecil namun kekuatan setiap pohon cukup besar. Hal tersebut ditunjukkan dengan diperolehnya forest dengan *error out of bag* bernilai kecil (Breiman, 2001). Dengan demikian, nilai *m* tersebut dapat diketahui melalui pengamatan *error out of bag* pada berbagai nilai *m*. Breiman (2001) menyarankan untuk mengamati *error out of bag* saat peubah penjelas, $m = \{1/2\ |\sqrt{p}|, |\sqrt{p}|, 2|\sqrt{p}|\}$ dan $k$ kecil, lalu memilih *m* yang menghasilkan *error out of bag* terkecil.

## 2.4 Evaluasi Ketepatan Klasifikasi

Evaluasi ketepatan klasifikasi yang digunakan pada penelitian ini adalah *total accuracy rate, sensitivity, specificity,* dan *area under curve* (AUC). *Total accuracy rate* merupakan proporsi

observasi yang diprediksi secara benar oleh fungsi klasifikasi. Sedangkan *sensitivity* merupakan ukuran proporsi yang benar-benar positif, sebaliknya *specificity* merupakan ukuran proporsi yang benar-benar negatif (Bekkar *et al*. 2013). *Area under curve* (AUC) merupakan sebuah indikator kinerja kurva ROC (*Receiver Operating Charateristic*) yang dapat meringkas kinerja *classifier* menjadi matriks single. Perhitungan ketepatan klasifikasi ditunjukkan pada tabulasi silang pada Tabel 2.1.

**Tabel 2.1** Tabulasi Silang Ketepatan Klasifikasi

| **Kelas Pengamatan Y** | **Kelas Prediksi Y** | | **Total** |
|---|---|---|---|
| | **1** | **2** | |
| **1** | $m_{11}$ | $m_{12}$ | $M_{1.}$ |
| **2** | $m_{21}$ | $m_{22}$ | $M_{2.}$ |
| **Total** | $M_{.1}$ | $M_{.2}$ | $M$ |

Keterangan:

$m_{ij}$ = jumlah pengamatan dari variabel Y kelas *i* yang tepat diprediksi sebagai variabel Y kelas *j*. ($i = j$)

$m_{ij}$ = jumlah pengamatan dari variabel Y kelas *i* yang salah diprediksi sebagai variabel Y kelas *j*. ($i \neq j$)

$M_{i.}$ = jumlah pengamatan dari variabel Y kelas *i*.

$M_{.j}$ = jumlah prediksi dari variabel Y kelas *j*.

$M$ = jumlah total pengamatan atau prediksi.

Berikut persamaan untuk menghitung *total accuracy rate, sensitivity, specificity,* dan *area under curve.*

$$total\ akurasi = \frac{M_{1.}}{M} \quad (2.4)$$

$$sensitivity = \frac{m_{11}}{M_{1.}} \quad (2.5)$$

$$spesificity = \frac{m_{22}}{M_{2.}} \tag{2.6}$$

$$AUC = \frac{1}{2}(sensitifity + spesificity) \tag{2.7}$$

Perhitungan AUC dilakukan melalui beberapa pendekatan, yang paling banyak digunakan adalah *trapezoidal method*. Pendekatan tersebut berbasis metode geometris yang berdasarkan interpolasi kinier antara masing-masing titik pada kurva ROC (Bekkar, *et al*., 2013). Kategori berdasarkan nilai AUC dapat disajikan ke dalam tabel berikut.

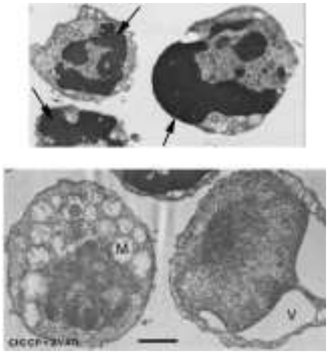
**Tabel 2.2** Kategori Kebaikan Model

| Nilai AUC | Kategori |
|---|---|
| 0,91-1,00 | *Excellent* (Sempurna) |
| 0,81-0,90 | *Very Good* (Sangat Baik) |
| 0,71-0,80 | *Good* (Baik) |
| 0,61-0,70 | *Fair* (Cukup) |
| 0,51-0,60 | *Poor* (Lemah) |

## 2.5 Kanker

Penyakit kanker merupakan suatu penyakit yang disebabkan pertumbuhan sel-sel jaringan tubuh tidak normal (tumbuh sangat cepat dan tidak terkendali), menginfiltrasi menekan jaringan tubuh sehingga mempengaruhi organ tubuh (Wahyuni, 2015). Penyakit kanker merupakan penyakit berat dan bersifat kronis, yang ditandai 19 pertumbuhan sel tubuh tidak normal, berkembang cepat, menyebar, dan menekan organ atau saraf sekitar. Pertumbuhan sel kanker tidak terkendali disebabkan kerusakan *deoxyribose nucleic acid* (DNA), sehingga menyebabkan mutasi gen vital yang mengontrol pembelahan sel. Beberapa mutasi dapat mengubah sel normal menjadi sel kanker. Mutasi-mutasi tersebut diakibatkan agen kimia maupun fisik yang edisebut karsinogen. Mutasi dapat

terjadi secara spontan maupun diwariskan (Wahyuni, 2015). Sel-sel kanker membentuk suatu masa dari jaringan ganas yang kemudian menyusup ke jaringan di dekatnya dan menyebar ke seluruh tubuh. Sel-sel kanker sebenarnya dibentuk dari sel normal melalui proses transformasi terdiri dari dua tahap yaitu tahap iniasi dan promosi. Tahap inisiasi, pada tahap ini perubahan bahan genetis sel yang memancing sel menjadi ganas. Perubahan sel genetis disebabkan unsur pemicu kanker yang terkandung dalam bahan kimia, virus, radiasi, atau sinar matahari.



**Gambar 2.3** Metastasis dan Apoptosis Gen p53

## 2.6 Toksisitas

Seiring dengan kemajuan teknologi di bidang medis, kanker bukan lagi menjadi momok yang menakutkan bagi penderitanya. Banyak penderita kanker yang sembuh dan berhasil melewati pengobatan kemoterapi, radiasi, dan operasi. Kemoterapi adalah pengobatan yang dilakukan dengan cara memberikan obat-obatan yang berfungsi untuk melawan serta menghancurkan pertumbuhan kanker. Dalam dunia medis, pengobatan ini dianggap memiliki kekuatan yang cukup bisa menghentikan pertumbuhan dan perkembangan sel kanker di dalam tubuh. Cara kerja obat kemoterapi yaitu dengan cara menghancurkan dan menghentikan sel-sel yang sedang tumbuh. Sel kanker yang tumbuh secara abnormal tentu saja akan dihancurkan oleh obat-obat tersebut. Tetapi, tidak hanya sel kanker saja sel normal yang sedang tumbuh

dan berkembang juga akan dihancurkan, karena obat kemoterapi tidak bisa membedakan sel kanker dan sel normal yang sedang tumbuh atau dinamakan dengan toksisitas terhadap sel normal. Akibat toksisitas, sel normal yang sedang tumbuh akan rusak dan kemudian menyebabkan gangguan pada jaringan yang sedang berkembang. Dengan semakin tinggi kadar toksisitas maka sel normal yang hancur juga semakin banyak. Jika hampir semua sel di dalam tubuh sedang berkembang, maka bisa jadi kemoterapi menyebabkan kerusakan di semua jaringan (Wirasuta, 2006)

## 2.7 Radioterapi

Menurut Kementerian Kesehatan RI (2015), faktor-faktor penyebab penyakit kanker antara lain faktor genetik, faktor karsinogen, dan faktor gaya hidup. Pengobatan kanker biasanya menggunakan radioterapi. Radioterapi ini dapat digunakan untuk mengobati kanker primer stadium awal hingga stadium lanjut dengan menggunakan sinar gamma atau sinar *X* dan partikel lainnya untuk mematikan sel-sel kanker tanpa akibat fatal pada jaringan sehat di sekitarnya. Prinsip radioterapi adalah memberikan dosis radiasi yang mematikan tumor pada daerah yang telah ditentukan (volume target) sedangkan jaringan normal sekitarnya mendapat dosis seminimal mungkin. Hal ini sangat ditunjang dengan kemajuan teknologi dari alat-alat radioterapi dan kemajuan dari komputer. Perkembangan teknologi di dunia kedokteran tidak dapat dipungkiri telah membantu penderita penyakit untuk sembuh dari sakit yang dideritanya dan meningkatkan kualitas hidup penderita tersebut.

Hingga saat ini, belum ada proteksi radiasi yang ideal untuk senyawa sintetik. Senyawa sintetik yang digunakan sebagai proteksi radiasi meliputi aminothiol, ethyl phosphorothioic acid yang dikenal juga sebagai amifostine, ethiophos, atau gammaphos. Senyawa sintetik tersebut menyebabkan efek toksik pada konsentrasi yang tertentu sehingga terdapat keterbatasan dalam penggunaannya. Menurut data *Food and Drug Administration*, yang diakui sebagai proteksi radiasi adalah amifostine. Radioterapi

memiliki beberapa efek bagi penggunanya, yaitu efek biologis, efek genetik, efek somatik, dan efek radiasi. Efek biologis yang terjadi berdasarkan jenis sel yaitu efek genetik dan efek somatik. Sedangkan efek genetik terjadi pada sel genetik yang akan diturunkan pada keturunan individu yang terpapar dan efek somatik akan diderita oleh individu yang terpapar radiasi. Efek radiasi ditinjau dari segi dosis radiasi dapat dibedakan berupa efek stokastik dan deterministik. Efek stokastik adalah efek akibat paparan sinar-X yang timbul setelah rentang waktu tertentu tanpa adanya batas ambang dosis sedangkan efek deterministik adalah efek yang langsung terjadi apabila paparan sinar-X melebihi ambang batas dosis dimana tingkat keparahan bergantung pada dosis radiasi yang diterima (CH, 2011).

*(Halaman ini sengaja dikosongkan)*

# BAB III
# METODOLOGI PENELITIAN

## 3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang berasal dari penelitian yang dilakukan oleh Matohashi *et al*. (2018) yaitu data perankingan senyawa yang berhubungan dengan gen supresor p53. Variabel prediktor pada penelitian ini berjumlah 217 variabel.

## 3.2 Variabel Penelitian

Variabel yang digunakan dalam penelitian ini terdiri dari dua variabel respon (Y) yaitu variabel respon untuk *cytotoxycity* dan *radioprotective* serta menggunakan 217 variabel prediktor. Kedua struktur data untuk variabel respon dipisahkan karena akan dilakukan pengujian yang dilakukan sendiri-sendiri (parsial) pada kedua variabel respon, sehingga hasil yang didapatkan tidak akan mempengaruhi fungsi masing-masing variabel. Berikut Tabel 3.1 dan Tabel 3.2 merupakan struktur data.

**Tabel 3.1** Struktur Data Penelitian untuk Toksisitas ($Y_1$)

| No | Kelas | Variabel Prediktor | | | | | |
|---|---|---|---|---|---|---|---|
| | | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $\cdots$ | $X_{217}$ |
| 1 | 1 | $X_{1,1}$ | $X_{2,1}$ | $X_{3,1}$ | $X_{4,1}$ | $\cdots$ | $X_{217,1}$ |
| 2 | 0 | $X_{1,2}$ | $X_{2,2}$ | $X_{3,2}$ | $X_{4,2}$ | $\cdots$ | $X_{217,2}$ |
| 3 | 0 | $X_{1,3}$ | $X_{2,3}$ | $X_{3,3}$ | $X_{4,3}$ | $\cdots$ | $X_{217,3}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| 20 | 0 | $X_{1,20}$ | $X_{2,20}$ | $X_{3,20}$ | $X_{4,20}$ | $\cdots$ | $X_{217,20}$ |

**Tabel 3.2** Struktur Data Penelitian untuk Proteksi Radiasi ($Y_2$)

| No | Kelas | Variabel Prediktor | | | | | |
|---|---|---|---|---|---|---|---|
| | | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $\cdots$ | $X_{217}$ |
| 1 | 0 | $X_{1,1}$ | $X_{2,1}$ | $X_{3,1}$ | $X_{4,1}$ | $\cdots$ | $X_{217,1}$ |
| 2 | 1 | $X_{1,2}$ | $X_{2,2}$ | $X_{3,2}$ | $X_{4,2}$ | $\cdots$ | $X_{217,2}$ |
| 3 | 1 | $X_{1,3}$ | $X_{2,3}$ | $X_{3,3}$ | $X_{4,3}$ | $\cdots$ | $X_{217,3}$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\ddots$ | $\vdots$ |
| 20 | 1 | $X_{1,20}$ | $X_{2,20}$ | $X_{3,20}$ | $X_{4,20}$ | $\cdots$ | $X_{217,20}$ |

Pada Tabel 3.1 dan Tabel 3.2, variabel prediktor yang digunakan berjumlah 217 prediktor. Variabel prediktor merupakan karakteristik senyawa obat kanker yang didapatkan pada penelitian sebelumnya melalui aplikasi *software Discovery Studio* (*software* pemodelan yang digunakan untuk menghitung sifat kimia). Struktur data untuk masing-masing variabel respon dipisahkan karena pada keduanya akan dilakukan klasifikasi yang terpisah untuk mendapatkan perankingan yang optimal untuk kadar toksisitas dan fungsi radio proteksi yang akan dihasilkan oleh senyawa gen supresor p53. Pada masing-masing variabel respon memiliki dua target kelas. Target kelas untuk toksisitas yaitu 1 untuk toksisitas dengan tingkat kematian sel normal 20-100% dan 0 untuk Toksisitas dengan tingkat kematian sel normal kurang dari 20%. Sedangkan target kelas untuk radio proteksi yaitu 1 untuk proteksi radiasi rendah dan 0 untuk proteksi radiasi tinggi. Berikut Tabel 3.3 merupakan variabel penelitian yang akan digunakan dalam penelitian ini

**Tabel 3.3** Variabel Penelitian

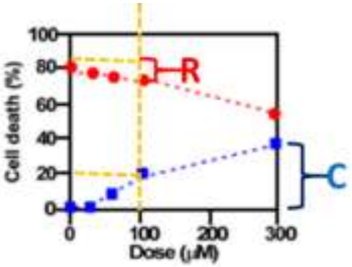
| Variabel | Nama Variabel | Skala |
|---|---|---|
| Respon ($Y_1$) | Target kelas<br>$Y(0)$ = Toksisitas dengan tingkat kematian sel normal 20-100%<br>$Y(1)$ = Toksisitas dengan ingkat kematian sel normal < 20% | Nominal |
| Respon ($Y_2$) | Target kelas<br>$Y(0)$ = Proteksi radiasi rendah (tingat kematian sel kanker ≤10 %)<br>$Y(1)$ = Proteksi radiasi tinggi (tingat kematian sel kanker >10 %) | Nominal |
| Prediktor ($x_i$) | $x_1$ = pKa (tingkat keasaman) | Rasio |
| | $x_2$ = Jumlah atom Br (Br_Count) | Rasio |
| | $x_3$ = Jumlah atom C (C_Count) | Rasio |
| | $x_4$ = Jumlah atom CI (Cl_Count) | Rasio |
| | $x_5$ = Jumlah atom F (F_Count) | Rasio |

**Tabel 3.3** Variabel Penelitian (lanjutan)

| | | |
|---|---|---|
| Prediktor ($x_i$) | $x_8$ = Jumlah atom N (N_Count) | Rasio |
| | ⋮ | |
| | $x_{212}$ = Luas area molekul di bidang YZ (Shadow_YZfrac) | Rasio |
| | $x_{213}$ = Panjang molekul pada dimensi Z (Shadow_Zlength) | Rasio |
| | $x_{214}$ = Jumlah molekul 3D Polar SASA (Molecular_3D_PolarSASA) | Rasio |
| | $x_{215}$ = Jumlah molekul 3D SASA (Molecular_3D_SASA) | Rasio |
| | $x_{216}$ = Jumlah molekul 3D SAVol (Molecular_3D_SAVoL) | Rasio |
| | $x_{217}$ = Volume molekul (Molecular_Volume) | Rasio |

Variabel respon pada Tabel 3.3 merupakan target kelas yang terdiri dari dua kelas (biner), yaitu 0 dan 1. Pada respon $Y_1$, jumlah respon kelas 0 sebanyak 18 dan kelas 1 sebanyak 2. Sedangkan untuk respon $Y_2$, jumlah respon kelas 0 sebanyak 5 dan kelas 1 sebanyak 15.

Klasifikasi pada $Y_1$ didasarkan pada tingkat kematian pada sel normal, kelas 1 untuk toksisitas dengan tingkat kematian sel normal antara 20 – 100 % dan kelas 0 untuk toksisitas dengan tingkat kematian sel normal antara >100 %. Untuk klasifikasi pada $Y_2$ didasarkan pada tingkat proteksi radiasi pada sel normal, kelas 1 untuk proteksi radiasi dengan tingkat kematian sel normal ≤10% dan kelas 0 untuk proteksi radiasi dengan tingkat kematian sel normal >10%. Batas untuk toksisitas dan radio proteksi tersebut didapatkan berdasarkan titik potong dari tingkat kematian sel (20%) dan dosis senyawa ($100\mu M$). Berikut ini adalah gambar pada penelitian Matohashi *et al* (2018) dalam menentukan *threshold* untuk kelas klasifikasi pada toksisitas pada Gambar 3.1.




Sumber : Matohashi *et al* (2018)

**Gambar 3.1** Penentuan *threshold* Pada Kelas *Cytotoxicity* dan *Radioprotective*

C merupakan *threshold* pada *cytotoxicity* dan R merupakan *threshold* pada *radioprotective*. Penentuan *threshold* pada toksisitas dilakukan dengan perhitungan persentase minimum tingkat kematian yaitu pada dosis $100\mu M$, sehingga didapatkan persentase maksimum tingkat kematian sel yaitu pada dosis $300\mu M$ dikurangi dengan *threshold* tingkat kematian sel untuk toksisitas yaitu 20%. Sedangkan untuk proteksi radiasi, penentuan *threshold* dengan perhitungan persentase maksimum tingkat kematian sel yaitu pada dosis $0\mu M$ dikurangi dengan persentase minimum tingkat kematian yaitu pada dosis $20\mu M$, sehingga didapatkan *threshold* tingkat kematian sel untuk proteksi radiasi yaitu 10%.

## 3.3 Langkah Analisis

Sebelum melakukan analisis dengan metode *neural network ensemble* dan *random forest*, terlebih dahulu dilakukan eksplorasi data untuk mengetahui karakteristiknya. Langkah-langkah analisis dalam penelitian ini adalah sebagai berikut.

1. Untuk mendapatkan hasil klasifikasi secara univariat terhadap senyawa gen supresor p53 pada tingkat toksisitas dan pada fungsi proteksi radiasi.
   a. Membuat analisa karakteristik terhadap data yaitu diagram lingkaran untuk melihat perbandingan kelas *cytotoxicity* dan *radioprotective* yang rendah dan tinggi serta diagram batang pada setiap variabel terhadap kelas.

b. Melakukan analisis klasifikasi menggunakan metode *Neural network Ensemble* dengan langkah sebagai berikut.
i. Menentukan jumlah *subspace* (k) dimana k = 12, 13, 14, 15, 16, 17, 18, dan 19. Penentuan jumlah *subspace* yang digunakan menggunakan persamaan sebagai beikut.

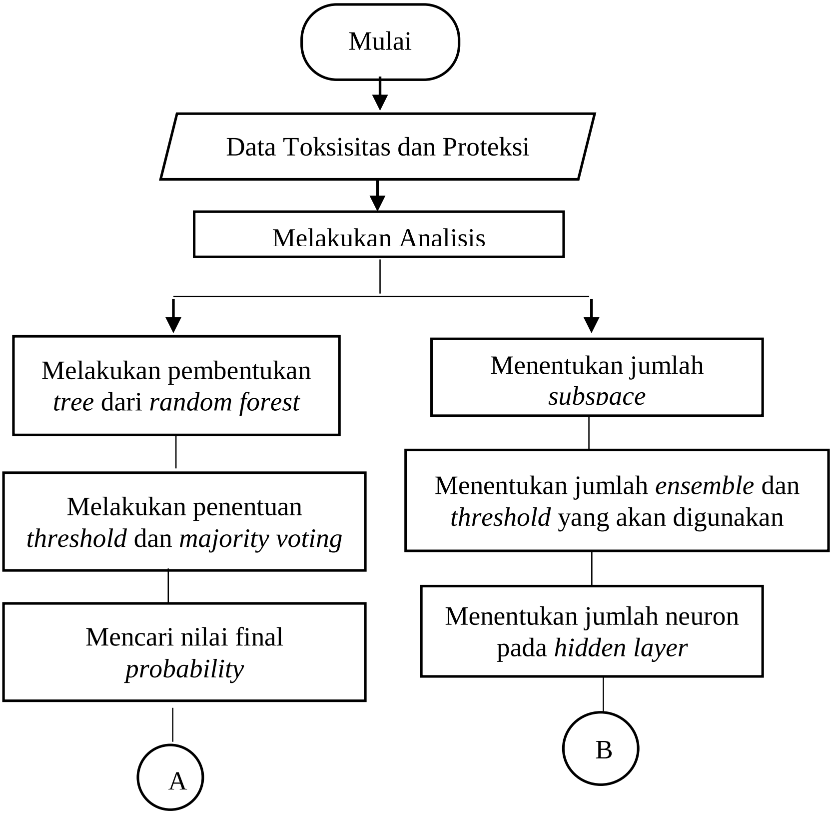
$$\frac{p}{n-1} \tag{3.1}$$

dimana p adalah variabel prediktor dan n merupakan jumlah observasi. Variabel yang berlebih pada *subspace* secara otomatis akan masuk kedalam partisi secara berurut kedalam *subspace* pertama, kedua, dan seterusnya.
ii. Menentukan banyak *ensemble* (j) dan nilai *threshold* yang akan digunakan. Jumlah *ensemble* yang digunakan adalah 11 dengan threshold 0.5 dan threshold optimum.
iii. Menentukan jumlah neuron pada hidden layer
iv. Melakukan estimasi parameter pada pembobot *neural network.*
v. Mendapatkan nilai probabilitas klasifikasi dari masing-masing *subspace* dan mengklasifikasikannya dengan menggunakan *threshold* (0.5 dan optimum) dan mengulangi langkah hingga terbentuk 11 *ensemble*.
vi. Membandingkan hasil nilai threshold 0.5 dengan threshold optimum.
c. Melakukan analisis klasifikasi menggunakan metode *random forest* dengan langkah sebagai berikut
i. Melakukan pembentukan *random forest* dengan membentuk 500 *trees* dan variabel prediktor yang dipecah sejumlah 72.
ii. Menentukan *threshold* untuk menentukan kelas prediksi dan kemudian dilakukan *majority voting* untuk penentuan kelas klasifikasi.
iii. Menghitung nilai rata-rata probabilitas dari kelas prediksi terbanyak pada hasil *majority voting* sebagai *final probability*.

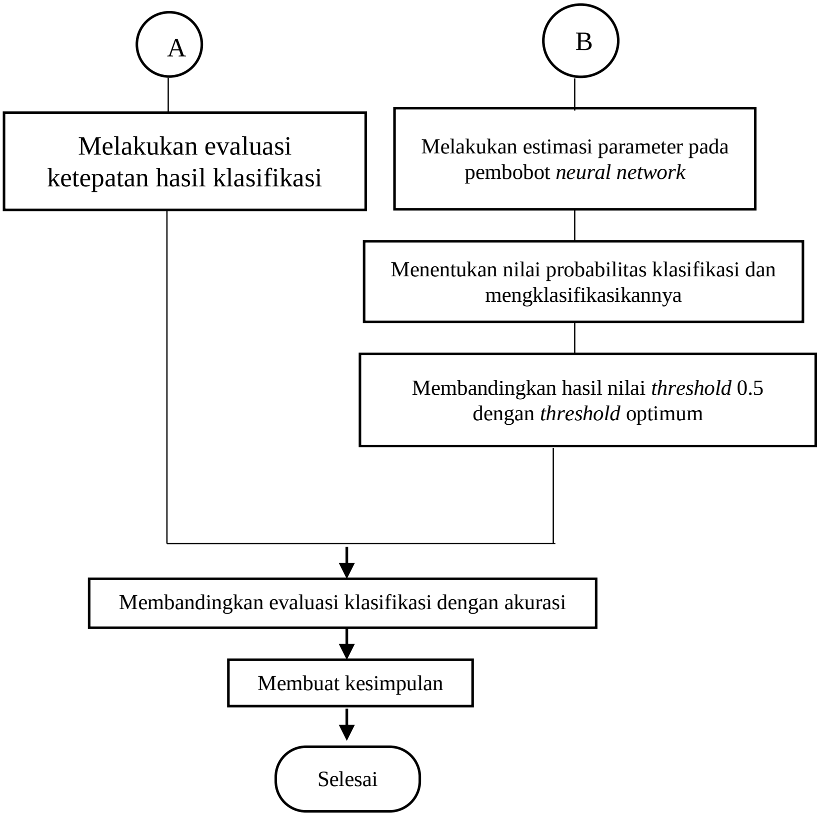
iv. Melakukan evaluasi ketepatan hasil klasifikasi dengan Melakukan evaluasi hasil klasifikasi dengan nilai *AUC* pada tiap kelas.

2. Menghitung akurasi total dari model klasifikasi *neural network ensemble* dan *random forest* yang terbentuk dan memilih metode terbaik berdasarkan akurasi terbaik.

Berdasarkan langkah analisis diatas, diagram alir analisis dalam penelitian ini secara umum dapat diilustrasikan seperti pada Gambar 3.2.




**Gambar 3.2** Diagram Alir Penelitian




**Gambar 3.2** Diagram Alir Penelitian (lanjutan)

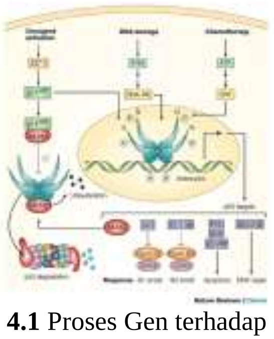
*(Halaman ini sengaja dikosongkan)*

# BAB IV
# ANALISIS DAN PEMBAHASAN

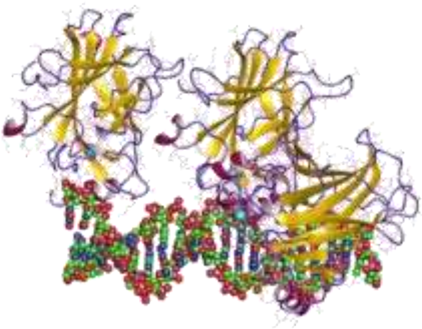
Pada penelitian ini dilakukan evaluasi klasifikasi dengan menggunakan *artificial neural network* ensemble dan *random forest.* Klasifikasi yang dilakukan berdasarkan dua kelas toksisitas dan radio proteksi guna mengetahui senyawa obat terbaik untuk kanker. Kemudian akan dilakukan perbandingan nilai akurasi dari setiap metode yang digunakan.

## 4.1 Karakteristik Data *Citotoxicity* dan *Radioprotective*

Sel kanker didefinisikan sebagai sel dengan proliferasi yang berlawanan dari kontrol normalnya. Selain itu, sel kanker mampu untuk menyerang dan berkolonisasi di jaringan sekitar. Suatu sel dikatakan menjadi sel kanker bila telah mengabaikan sinyal eksternal dan internal yang meregulasi proliferasi sel, dapat menghindari apoptosis, secara genetik tidak stabil, bisa keluar dari jaringan awalnya (invasif), bisa bertahan hidup dan berproliferasi di jaringan lain. Banyak hal dapat menyebabkan terjadinya kanker, diantaranya adalah terjadinya kerusakan DNA yang berakhir dengan mutasi gen. Mutagen yang bisa mengakibatkan kanker antara lain bahan kimia karsinogenik, virus, berbagai bentuk radiasi termasuk sinar UV, dan radiasi ionisasi seperti sinar gamma dan partikel alpha yang keluar dari radioaktif. Illustrasi berikut menjelaskan detail apabila terdapat gen yang rusak akibat kanker.



**Gambar 4.1** Proses Gen terhadap Mutagen



**Gambar 4.2** Gen Supresor p53

Pengobatan kanker yang selama ini dilakukan bukan saja berdampak pada sel kanker namun juga berdampak pada sel normal yang berada disekitarnya. Sehingga dibutuhkan adanya pengobatan kanker dengan kadar toksisitas minimal dan proteksi radiasi yang maksimal. Berdasarkan data *citotoxicity* dan *radioprotective*, penelitian ini mendapatkan karateristik kedua tersebut yang secara deskriptif dapat dijelaskan dengan Gambar 4.3 di bawah ini.

**Gambar 4.3** Proporsi Jumlah Kelas Respon Data *Citotoxicity* dan *Radioprotective*

Pada Gambar 4.3 menunjukkan bahwa komponen senyawa gen p53 memiliki dua kelas pada respon dengan presentase pada masing-masing kelas yaitu sebesar 10% atau sebanyak dua senyawa untuk kelas 1 dan 90% atau sebanyak 18 senyawa untuk kelas 0. Untuk data *radioptotective* diketahui bahwa presentase kelas 1 sebesar 25 % atau sebanyak 5 senyawa dan presentase kelas 0 sebesar 75 % atau sebesar 15 senyawa. Baik pada data *citotoxicity* maupun radioprotection. keduanya memiliki proporsi yang tidak seimbang atau *imbalance*, sehingga penggunaan batas *threshold* 0.5 kurang tepat dan diperlukan perhitungan *threshold* optimal dalam proses klasifikasi yang akan dilakukan (Lim, Ahn, Moon, & Chen, 2009). Klasifikasi dilakukan dengan menggunakan 217 variabel prediktor yang merupakan karakteristik dari komponen senyawa tersebut.

## 4.2 Klasifikasi dengan *ANN Ensemble*

Untuk memudahkan pembacaan alur proses *artificial neural network ensemble*, maka diberikan ilustrasi proses metode *artificial neural network* sebagai berikut.

### 4.2.1 Ilustrasi Metode *ANN Ensemble*

Bagian ini menjelaskan bagaimana *artificial neural network ensemble* untuk melakukan klasifikasi terhadap data *citotoxicity* dan *radioprotective*. Berikut diberikan illustrasi mengenai langkah melakukan klasifikasi menggunakan metode ANN *Ensemble* untuk data *citotoxicity* dengan fungsi aktifasi *softmax* pada 12 *subspace*.

1. Dengan menggunakan *all training set* dengan jumlah variabel prediktor ($p$) yang digunakan sebanyak 217 dan jumlah data ($n$) sebanyak 20, maka data yang digunakan tergolong dalam data berdimensi tinggi karena jumlah prediktor lebih banyak dari jumlah pengamatan. Hal ini bisa ditulis dengan $p > n$.
2. *Node-node* pada *input layer* memiliki jumlah *subspace* atau partisi yang akan digunakan yaitu 12 *subspace*, 13 *subspace*, 14 *subspace*, 15 *subspace*, 16 *subspace*, 17 *subspace*, 18 *subspace*, 19 *subspace*. Jumlah ini mempertimbangan dari perhitungan jumlah minimal *subspace* atau partisi yang bisa digunakan. Dengan menggunakan keseluruhan data dimana $n$ = 20 dan $p$ =

217, maka jumlah maksimal *subspace* yang bisa digunakan dalam satu *subspace* atau partisi adalah $n - 1 = 20 - 1 = 19$. Jumlah *subspace* atau partisi ($m$) minimal yang bisa digunakan agar memenuhi syarat $p < n$ maka diperoleh hasil sebagai berikut

$$m = \frac{p}{n-1} = \frac{217}{20-1} = 11{,}42 \approx 12 \tag{4.1}$$

Sehingga jumlah *subspace* atau partisi minimal yang terbentuk adalah berjumlah 12. Dengan *subspace* minimal yang terbentuk adalah 12, maka banyak partisi yang akan digunakan berjumlah 8 partisi yang terdiri dari 12 *subspace* , 13 *subspace* , 14 *subspace* , 15 *subspace* , 16 *subspace* , 17 *subspace* , 18 *subspace* , dan 19 *subspace*.

3. Kemudian *node-node* pada *input layer* disalurkan ke *node-node hidden layer* dan dilakukan pengestimasian nilai bobot dan bias. Setelah didapatkan nilai estimasi bobot dan bias, *node-node* pada *hidden layer* dikalikan dengan nilai bobot dan dijumlahkan dengan nilai bias sesuai dengan persamaan berikut.

$$v_j^h = \sum_{i=1}^{p} w_{ji}^h X_i + b_j^h \tag{4.2}$$

dengan,

$X_i$ = variabel *input* sebanyak $p$, ($i = 1,2,\ldots,p$)

$w_{ji}^h$ = bobot dari *input* ke-$i$ yang menuju *neuron* ke-$j$ pada *hidden layer*, dengan $j = 1,2,\ldots,q$

$b_j^h$ = bias pada *neuron* ke-$j$ pada *hidden layer*, ($j = 1,2,\ldots,q$)

Nilai $v_j^h$ merupakan suatu vektor nilai-nilai setelah proses penjumlahan input dan bobot.

4. Setelah didapatkan penjumlahan dengan nilai bias dan bobot, lalu menghitung nilai *output* pada *hidden layer* yang akan

terproses di *neuron* ke-j dengan fungsi aktivasi yang akan digunakan dengan menggunakan persamaan sebagai berikut.

$$a_j^h = f_j^h(v_j^h), \quad (4.3)$$

Dimana $f_j^h$ adalah fungsi aktifasi yang akan digunakan. Dengan menggunakan persamaan di atas, maka untuk data ke-t didapatkan persamaan berikut.

$$a_{j(t)}^h = f_j^h\left(\sum_{i=1}^{p} w_{ji}^h X_{i(t)} + b_j^h\right). \quad (4.4)$$

5. Menentukan *threshold* pada masing-masing data untuk menentukan kelas prediksi yang didapatkan pada *output layer*. Berikut *threshold* untuk *citotoxicity*

$$threshold = \frac{\frac{1}{20}\sum_{i=1}^{20} y_i + 0,5}{2} = \frac{0,1+0,5}{2} = 0.3 \quad (4.5)$$

Dan *threshold* untuk *radioprotective*

$$threshold = \frac{\frac{1}{20}\sum_{i=1}^{20} y_i + 0,5}{2} = \frac{0,75+0,5}{2} = 0,625 \quad (4.6)$$

Sehingga nilai *threshold* yang digunakan untuk data *citotoxicity* adalah 0.3 dan untuk data *radioprotective* yaitu 0.625. *Threshold* tersebut digunakan untuk menentukan prediksi target. Apabila nilai probabilitas prediksi $\geq$0,3 untuk *citotoxicity* dan nilai probabilitas prediksi $\geq$0,625 untuk *radioprotective*, maka akan masuk kelas 1 dan sebaliknya masuk kelas 0.

6. Analisis untuk *ensemble* pertama dilakukan dengan menggunakan fungsi aktivasi *softmax* tertera pada langkah berikut.

i. Sebanyak 217 variabel prediktor dibagi menjadi 12 *subspace* secara random sehingga pada masing-masing *subspace* akan memiliki sejumlah 18 variabel prediktor.

ii. Menghitung nilai probabilitas untuk masing-masing pengamatan setiap *subspace* berdasarkan persamaan 2.1.

Berikut merupakan probabilitas dan klasifikasi dari *ensemble* pertama.

**Tabel 4.1** Probabilitas *Ensemble* Pertama data *citotoxicity* dengan Fungsi Aktifasi *Softmax* Pada *Subspace* 12

| Obs | Probabilitas Ensemble 1 | Kelas |
|---|---|---|
| 1 | 0,3543 | 1 |
| 2 | 0,1806 | 0 |
| 3 | 0,0833 | 0 |
| 4 | 0,0816 | 0 |
| 5 | 0,0830 | 0 |
| 6 | 0,0849 | 0 |
| 7 | 0,1170 | 0 |
| 8 | 0,0977 | 0 |
| 9 | 0,0788 | 0 |
| 10 | 0,1452 | 0 |
| 11 | 0,1027 | 0 |
| 12 | 0,0932 | 0 |
| 13 | 0,0758 | 0 |
| 14 | 0,2251 | 0 |
| 15 | 0,0731 | 0 |
| 16 | 0,0680 | 0 |
| 17 | 0,0730 | 0 |
| 18 | 0,0653 | 0 |
| 19 | 0,0955 | 0 |
| 20 | 0,0768 | 0 |

Probabilitas pada tabel 4.1 diperoleh dari hasil substitusi nilai parameter pada tabel langkah ii ke dalam persamaan fungsi aktivasi yaitu softmax. Berikut penjabaran persamaan 2.1 dengan fungsi aktivasi *softamax* dimana merupakan persamaan yang digunakan untuk mendapatkan nilai probabilitas pada tabel 4.1.

$$\hat{Y}_{(t)} = \frac{\exp(x_i)}{\sum_{k=1}^{m}\exp(x_k)}\left[\sum_{j=1}^{q}\left[w_j^o\left(\frac{\exp(x_i)}{\sum_{k=1}^{m}\exp(x_k)}\right)\left(\sum_{i=1}^{p}w_{ji}^h X_{i(t)} + b_j^h\right) + b^o\right]\right] \quad (4.7)$$

dengan,
$X_{i(k)}$ = variabel *input* sebanyak *p*, (*i* = 1,2,…,*p*)
fungsi aktivasi yang digunakan sebagai berikut,

$$f(x_i) = \frac{\exp(x_i)}{\sum_{k=1}^{p} \exp(x_k)} \text{ , i= 1, 2, 3, ..., p}$$

iii. Mengulangi kedua langkah analisis di atas hingga terbentuk 11 *ensemble*. Pemilihan 11 ensemble digunakan untuk menghindari adanya nilai yang sama pada saat dilakukan majorirty voting. Selanjutnya. untuk menentukan prediksi *artificial neural network ensemble* paling akhir digunakan gabungan dari kesebelas prediksi *ensemble* dengan cara rata-rata. Berikut merupakan hasil keseluruhan *artificial neural network ensemble* dengan 11 *ensemble* pada 12 *subspace* dengan menggunakan fungsi aktivasi *softmax*.

**Tabel 4.2** Hasil Klasifikasi ANN pada *Citotoxicity* dengan 11 *Ensemble* Untuk 12 *Subspace* Menggunakan Aktifasi *Softmax*

| **Obs** | **Kelas *Ensemble* ke-** | | | | | |
|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** |
| 1 | 0.3543(1) | 0.3871(1) | 0.3606(1) | 0.3176(1) | 0.3433(1) | 0.2525(0) |
| 2 | 0.1806(0) | 0.1713(0) | 0.1838(0) | 0.1374(0) | 0.1294(0) | 0.2033(0) |
| 3 | 0.0833(0) | 0.0811(0) | 0.0990(0) | 0.0903(0) | 0.0751(0) | 0.0767(0) |
| 4 | 0.0816(0) | 0.0708(0) | 0.0990(0) | 0.0859(0) | 0.1197(0) | 0.0658(0) |
| 5 | 0.0830(0) | 0.0754(0) | 0.0864(0) | 0.0644(0) | 0.0732(0) | 0.0785(0) |
| 6 | 0.0849(0) | 0.0740(0) | 0.0948(0) | 0.0808(0) | 0.0703(0) | 0.0945(0) |
| 7 | 0.1170(0) | 0.0941(0) | 0.0811(0) | 0.0973(0) | 0.1018(0) | 0.0836(0) |
| 8 | 0.0977(0) | 0.0692(0) | 0.0922(0) | 0.0853(0) | 0.0736(0) | 0.0872(0) |
| 9 | 0.0788(0) | 0.0652(0) | 0.0812(0) | 0.0560(0) | 0.1014(0) | 0.0826(0) |
| 10 | 0.1452(0) | 0.1419(0) | 0.1483(0) | 0.1298(0) | 0.1620(0) | 0.1243(0) |
| 11 | 0.1027(0) | 0.0663(0) | 0.0899(0) | 0.0696(0) | 0.0688(0) | 0.0671(0) |
| 12 | 0.0932(0) | 0.0896(0) | 0.0992(0) | 0.1397(0) | 0.0768(0) | 0.0842(0) |
| 13 | 0.0758(0) | 0.0644(0) | 0.0455(0) | 0.0745(0) | 0.0608(0) | 0.0637(0) |
| 14 | 0.2251(0) | 0.3106(1) | 0.2687(0) | 0.2880(0) | 0.2745(0) | 0.3648(1) |
| 15 | 0.0731(0) | 0.0922(0) | 0.0980(0) | 0.0930(0) | 0.1048(0) | 0.0913(0) |
| 16 | 0.0680(0) | 0.0812(0) | 0.0628(0) | 0.0659(0) | 0.0796(0) | 0.0798(0) |
| 17 | 0.0730(0) | 0.0550(0) | 0.0663(0) | 0.0681(0) | 0.0799(0) | 0.0729(0) |

**Tabel 4.2** Hasil Klasifikasi ANN pada *Citotoxicity* dengan 11 *Ensemble* Untuk 12 *Subspace* Menggunakan Aktifasi *Softmax* (lanjutan)

| Obs | Kelas *Ensemble* ke- | | | | | |
|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** |
| 18 | 0.0653(0) | 0.0623(0) | 0.0629(0) | 0.0514(0) | 0.0556(0) | 0.0570(0) |
| 19 | 0.0955(0) | 0.0808(0) | 0.0806(0) | 0.0753(0) | 0.0713(0) | 0.0946(0) |
| 20 | 0.0768(0) | 0.0649(0) | 0.0914(0) | 0.0774(0) | 0.0666(0) | 0.0797(0) |

| Obs | Kelas *Ensemble* ke- | | | | | Prediksi Kelas |
|---|---|---|---|---|---|---|
| | **7** | **8** | **9** | **10** | **11** | |
| 1 | 0.3538(1) | 0.3555(1) | 0.4142(1) | 0.4239(1) | 0.3737(1) | 1 |
| 2 | 0.1923(0) | 0.1379(0) | 0.1976(0) | 0.1548(0) | 0.1382(0) | 0 |
| 3 | 0.0935(0) | 0.0813(0) | 0.0556(0) | 0.1051(0) | 0.0963(0) | 0 |
| 4 | 0.0857(0) | 0.0796(0) | 0.0678(0) | 0.0622(0) | 0.0857(0) | 0 |
| 5 | 0.0672(0) | 0.0753(0) | 0.0573(0) | 0.0840(0) | 0.0783(0) | 0 |
| 6 | 0.0948(0) | 0.0949(0) | 0.0970(0) | 0.0773(0) | 0.0859(0) | 0 |
| 7 | 0.1089(0) | 0.1129(0) | 0.1023(0) | 0.1024(0) | 0.0780(0) | 0 |
| 8 | 0.0743(0) | 0.0854(0) | 0.0585(0) | 0.0655(0) | 0.0624(0) | 0 |
| 9 | 0.0684(0) | 0.0774(0) | 0.0864(0) | 0.0554(0) | 0.0732(0) | 0 |
| 10 | 0.1551(0) | 0.1565(0) | 0.1719(0) | 0.1442(0) | 0.1586(0) | 0 |
| 11 | 0.0706(0) | 0.0782(0) | 0.0724(0) | 0.0780(0) | 0.0880(0) | 0 |
| 12 | 0.1082(0) | 0.1107(0) | 0.1079(0) | 0.1338(0) | 0.0880(0) | 0 |
| 13 | 0.0437(0) | 0.0687(0) | 0.0790(0) | 0.0437(0) | 0.0656(0) | 0 |
| 14 | 0.3780(1) | 0.2396(0) | 0.3361(1) | 0.3019(1) | 0.2984(0) | 0 |
| 15 | 0.0810(0) | 0.0959(0) | 0.0754(0) | 0.0790(0) | 0.1085(0) | 0 |
| 16 | 0.0720(0) | 0.0677(0) | 0.0595(0) | 0.0678(0) | 0.0745(0) | 0 |
| 17 | 0.0650(0) | 0.0978(0) | 0.0789(0) | 0.0580(0) | 0.0664(0) | 0 |
| 18 | 0.0466(0) | 0.0727(0) | 0.0511(0) | 0.0478(0) | 0.0582(0) | 0 |
| 19 | 0.0901(0) | 0.0650(0) | 0.0801(0) | 0.0982(0) | 0.0703(0) | 0 |
| 20 | 0.0864(0) | 0.0838(0) | 0.0618(0) | 0.0538(0) | 0.0924(0) | 0 |

Dari tabel di atas, nilai di dalam kurung merupakan prediksi kelas pada observasi di tiap kelas *ensemble* tersebut. Terlihat bahwa untuk observasi pertama pada setiap kelas *ensemble* didapatkan prediksi kelas 1 sebanyak 10 dan prediksi kelas 0 sebanyak 1. Dengan menggunakan *majority voting* diperoleh prediksi kelas untuk observasi pertama yaitu 1 karena jumlah kelas *ensemble* pada observasi pertama dengan prediksi kelas 1 lebih banyak daripada prediksi kelas 0. Hal yang sama juga dilakukan pada observasi kedua hingga observasi terakhir. Maka didapatkan hasil untuk observasi kedua hingga terakhir

memiliki prediksi kelas 0 karena jumlah prediksi kelas 0 lebih banyak daripada prediksi kelas 1 pada tiap kelas ensemblenya. Nilai-nilai pada tabel 4.2 didapatkan melalui persamaan 4.7

**7.** Evaluasi ketepatan yang digunakan adalah *sensitifity*. *specificity*. dan akurasi dengan rumus pada persamaan (2.4) hingga persamaan (2.7). Berikut tabulasi silang dari hasil klasifikasi pada data ilustrasi.

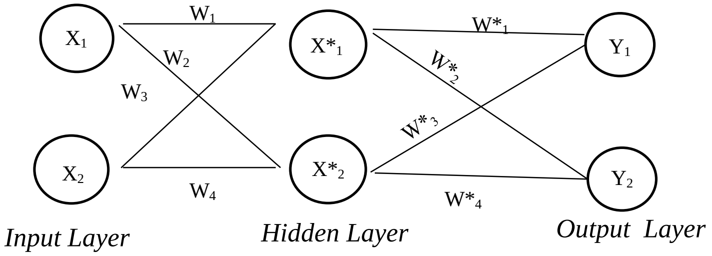
**Tabel 4.3** Tabulasi Silang Data Ilustrasi

| **Klasifikasi** | | | **Kelas Aktual** | |
|---|---|---|---|---|
| | | | **Negatif** | **Positif** |
| **Kelas Prediksi** | **12 *Subspace*** | **Negatif** | 18 | 1 |
| | | **Positif** | 0 | 1 |

Dari tabel tabulasi silang diperoleh hasil bahwa terdapat 18 pengamatan yang tepat terklasifikasi pada kelas negatif atau kelas 0 dan teradapat 1 pengamatan yang tepat terklasifikasi pada kelas positif atau kelas 1. Sehingga akurasi yang didapatkan untuk hasil tabulasi silang di atas sebagai berikut

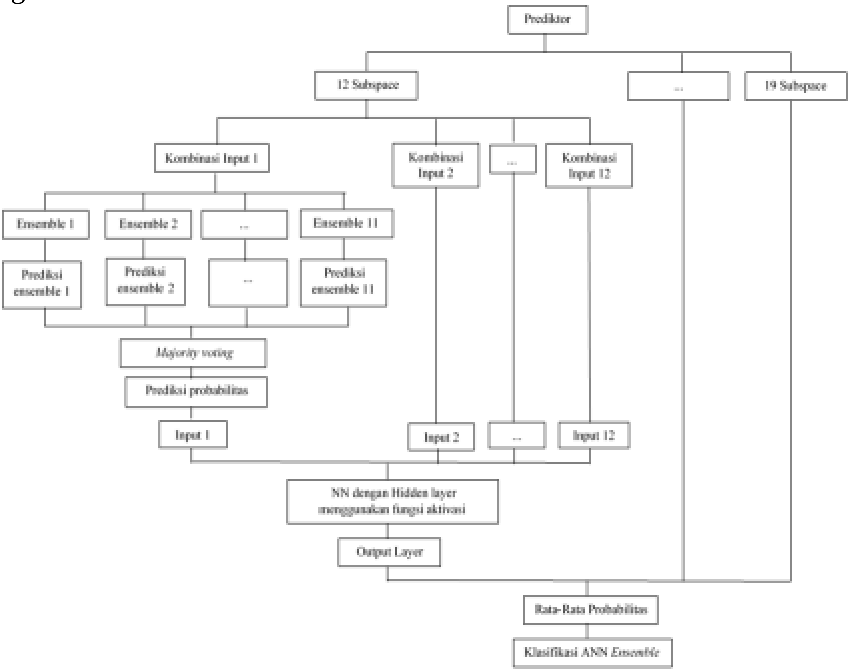
$$akurasi = \frac{18+1}{18+1+0+1} = \frac{19}{20} = 0,95 \qquad (4.7)$$

Maka untuk *ensemble* pertama memiliki nilai akurasi sebesar 95%. Bila langkah-langkah di atas digambarkan dalam bentuk sederhana, maka akan terlihat dalam bagan sebagai berikut.




**Gambar 4.6** llustrasi *Artificial Neural network Ensemble*

Dengan proses yang terjadi pada *input layer* seperti pada bagan berikut.




**Gambar 4.7** llustrasi Proses *Ensemble* Pada *Input Layer*

Melalui gambar 4.7 terlihat bila proses *ensemble* terjadi hanya pada *input layer* tanpa menggunakan fungsi aktivasi didalamnya. Setelah data terolah pada *input layer*, kemudian data memasuki *hidden layer* dan berakhir pada *output layer*. Pada kedua layer tersebut fungsi aktivasi mulai digunakan.

### 4.2.2 Klasifikasi Senyawa Penyusun Obat Kanker Menggunakan ANN *Ensemble*

Analisis *artificial neural network ensemble* merupakan pendekatan komputasional yang dapat digunakan untuk menyelesaikan masalah klasifikasi. *Artificial neural network* tidak memiliki asumsi apapun untuk dipenuhi. Demi mendapatkan hasil

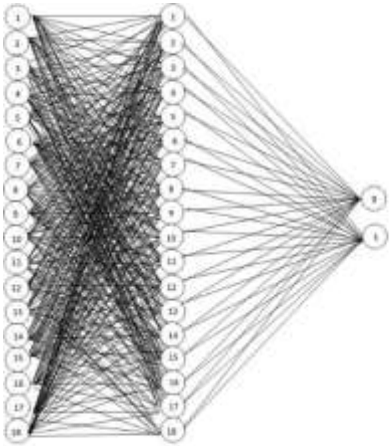
klasifikasi terbaik. analisis *artificial neural network ensemble* dilakukan beberapa kali dengan jumlah partisi (*subspace* ) yang berbeda pula. Dalam penelitian ini. partisi yang dibentuk adalah sebanyak 12, 13, 14, 15, 16, 17, 18, dan 19 *subspace* dengan menggunakan *threshold* optimal seperti perhitungan sebelumnya yaitu 0.3 untuk *citotoxicity* dan 0.625 untuk *radioprotective*. Nilai *ensemble* yang akan digunakan dalam penelitian ini adalah sebesar 11. nilai tersebut digunakan atas dasar agar tidak terjadi seri pada saat dilakukan *majority voting.*

Jumlah *node input* pada setiap *subspace* memiliki jumlah yang berbeda-beda. Hal tersebut dikarenakan jumlah *node-node* tersebut diperoleh dari jumlah variabel prediktor yang dibagi dengan jumlah *subspace* nya. Sedangkan untuk jumlah *node* pada *hidden layer* ditetapkan berjumlah sama dengan jumlah *node input*. Hal ini sesuai yang dijelaskan Heaton (2015) bahwa salah satu cara penentuan jumlah *node hidden layer* yaitu diantara jumlah *node output* dan *node input* ($n_{output} \leq n_{hidden} \leq n_{input}$).

**Tabel 4.4** Arsitektur ANN *Ensemble*

| **Jumlah *Subspace*** | **Node Input** | **Node Hidden Layer** | **Node Output** |
|---|---|---|---|
| 12 *Subspace* | 18 | 18 | |
| 13 *Subspace* | 17 | 17 | |
| 14 *Subspace* | 16 | 16 | |
| 15 *Subspace* | 14 | 14 | |
| 16 *Subspace* | 14 | 14 | 2 |
| 17 *Subspace* | 13 | 13 | |
| 18 *Subspace* | 12 | 12 | |
| 19 *Subspace* | 11 | 11 | |

Dalam penelitian ini digunakan dua fungsi aktivasi yaitu *Softmax Activation Function* dan *Sigmoid Activation Function.* Untuk memperjelas tabel 4.4, maka berikut merupakan illustrasi ANN dengan menggunakan 12 *subspace* yang digambarkan pada bagan berikut.



**Gambar 4.8** Illustrasi Pada ANN dengan 12 *Subspace*

**A.** ***Citotoxicity***

**a.** ***Softmax Activation Function***

Fungsi aktivasi berfungsi sebagai penentu suatu neuron harus aktif atau tidak berdasarkan *weighted sum* dari input. Selain menggunakan *threshold*. untuk memperkirakan probabilitas bersyarat pada tiap kelas maka digunakan fungsi logistik berganda atau biasa disebut sebagai fungsi *softmax* dengan persamaan berikut.

$$f(x_i) = \frac{\exp(x_i)}{\sum_{k=1}^{p} \exp(x_k)} \quad (4.8)$$

Hasil analisis *artificial neural network ensemble* dengan fungsi aktivasi *softmax* dengan menggunakan *threshold* optimal yang dibutuhkan untuk menyeimbangkan *sensitifity* dan *specificity* sesuai dengan persamaan 4.5 yaitu dengan menggunakan *threshold* optimal 0.3. Sehingga didapatkan hasil perhitungan nilai probabilitas setiap *subsapce* yang dapat dilihat pada tabel 4.5.

**Tabel 4.5** *Probability* Setiap *Subspace* untuk Fungsi Aktivasi *Softmax*

| **Obs.** | **Probabilitas Tiap *Subspace*** | | | | **Kelas** |
|---|---|---|---|---|---|
| | **12** | **13** | **14** | **15** | |
| 1 | 0.3590(1) | 0.3453(1) | 0.3215(1) | 0.3177(1) | 1 |
| 2 | 0.1692(0) | 0.1642(0) | 0.1595(0) | 0.1647(0) | 0 |
| 3 | 0.0965(0) | 0.0924(0) | 0.0971(0) | 0.0999(0) | 0 |
| 4 | 0.0817(0) | 0.0882(0) | 0.0831(0) | 0.0842(0) | 0 |
| 5 | 0.0767(0) | 0.0811(0) | 0.0866(0) | 0.0809(0) | 0 |
| 6 | 0.0858(0) | 0.0867(0) | 0.0887(0) | 0.0943(0) | 0 |
| 7 | 0.0951(0) | 0.0951(0) | 0.1071(0) | 0.1022(0) | 0 |
| 8 | 0.0775(0) | 0.0850(0) | 0.0860(0) | 0.0919(0) | 0 |
| 9 | 0.0723(0) | 0.0720(0) | 0.0744(0) | 0.0770(0) | 0 |
| 10 | 0.1524(0) | 0.1520(0) | 0.1549(0) | 0.1625(0) | 0 |
| 11 | 0.0694(0) | 0.0749(0) | 0.0867(0) | 0.0806(0) | 0 |
| 12 | 0.1004(0) | 0.1043(0) | 0.1077(0) | 0.1013(0) | 0 |
| 13 | 0.0595(0) | 0.0659(0) | 0.0615(0) | 0.0660(0) | 0 |
| 14 | 0.2731(0) | 0.2799(0) | 0.2519(0) | 0.2570(0) | 0 |
| 15 | 0.0905(0) | 0.1011(0) | 0.1097(0) | 0.1075(0) | 0 |
| 16 | 0.0788(0) | 0.0720(0) | 0.0784(0) | 0.0809(0) | 0 |
| 17 | 0.0730(0) | 0.0773(0) | 0.0814(0) | 0.0799(0) | 0 |
| 18 | 0.0652(0) | 0.0639(0) | 0.0685(0) | 0.0669(0) | 0 |
| 19 | 0.0766(0) | 0.0823(0) | 0.0799(0) | 0.0839(0) | 0 |
| 20 | 0.0879(0) | 0.0894(0) | 0.0868(0) | 0.0928(0) | 0 |
| **Obs.** | **Probabilitas Tiap *Subspace*** | | | | **Kelas** |
| | **16** | **17** | **18** | **19** | |
| 1 | 0.2978(0) | 0.2690(0) | 0.2711(0) | 0.2540(0) | 1 |
| 2 | 0.1641(0) | 0.1617(0) | 0.1596(0) | 0.1464(0) | 0 |
| 3 | 0.1049(0) | 0.1020(0) | 0.1037(0) | 0.1091(0) | 0 |
| 4 | 0.0868(0) | 0.0900(0) | 0.0979(0) | 0.0896(0) | 0 |
| 5 | 0.0859(0) | 0.0896(0) | 0.0964(0) | 0.0947(0) | 0 |
| 6 | 0.0917(0) | 0.0973(0) | 0.0984(0) | 0.0981(0) | 0 |
| 7 | 0.1037(0) | 0.1077(0) | 0.1084(0) | 0.1102(0) | 0 |
| 8 | 0.0866(0) | 0.0919(0) | 0.0884(0) | 0.0952(0) | 0 |
| 9 | 0.0832(0) | 0.0860(0) | 0.0908(0) | 0.0882(0) | 0 |
| 10 | 0.1554(0) | 0.1595(0) | 0.1482(0) | 0.1562(0) | 0 |
| 11 | 0.0856(0) | 0.0909(0) | 0.0909(0) | 0.0933(0) | 0 |
| 12 | 0.1023(0) | 0.1105(0) | 0.1100(0) | 0.1112(0) | 0 |
| 13 | 0.0756(0) | 0.0772(0) | 0.0779(0) | 0.0820(0) | 0 |
| 14 | 0.2529(0) | 0.2272(0) | 0.2266(0) | 0.2104(0) | 0 |
| 15 | 0.1095(0) | 0.1074(0) | 0.1122(0) | 0.1118(0) | 0 |
| 16 | 0.0857(0) | 0.0860(0) | 0.0919(0) | 0.0913(0) | 0 |
| 17 | 0.0914(0) | 0.0957(0) | 0.0941(0) | 0.0928(0) | 0 |
| 18 | 0.0726(0) | 0.0722(0) | 0.0810(0) | 0.0927(0) | 0 |

**Tabel 4.5** *Probability* Setiap *Subspace* untuk Fungsi Aktivasi *Softmax* (lanjutan)

| **Obs.** | **Probabilitas Tiap *Subspace*** | | | | **Kelas** |
|---|---|---|---|---|---|
| | **16** | **17** | **18** | **19** | |
| 19 | 0.0864(0) | 0.0974(0) | 0.1002(0) | 0.0950(0) | 0 |
| 20 | 0.0981(0) | 0.0970(0) | 0.0948(0) | 0.0988(0) | 0 |

Nilai di dalam kurung merupakan klasifikasi kelas pada masing-masing *subspace* yang telah diensemble sebanyak 11 kali pada tiap *subspace*. Kemudian nilai kelas prediksi didapatkan melalui majority voting dari nilai kelas prediksi tiap *subspace*. Nilai-nilai di atas didapatkan dari persamaan sebagai berikut.

$$f(x_i) = \frac{\exp(x_i)}{\sum_{k=1}^{p} \exp(x_k)} \tag{4.8}$$

p = banyak jumlah node pada setiap *subspace*
Sehingga persamaan ANN untuk fungsi *softmax* terbentuk sebagai berikut.

$$\hat{Y}_1(k) = \frac{w_1^* x_1^* + w_3^* x_2^*}{w_1^* x_1^* + w_2^* x_2^* + w_3^* x_3^* + w_4^* x_4^*} \tag{4.9}$$

$$\hat{Y}_2(k) = \frac{w_2^* x_1^* + w_4^* x_2^*}{w_1^* x_1^* + w_2^* x_2^* + w_3^* x_3^* + w_4^* x_4^*} \tag{4.10}$$

Dimana,
w* = *weight* pada *output layer*
x* = *node input* yang telah diolah pada *hidden layer*

Pada tabel 4.5 nilai di dalam tanda kurung merupakan prediksi kelas pada tiap *subpsace*. Selain itu pada tabel 4.5 juga terlihat bahwa hanya observasi 1 yang memiliki kadar toksisitas yang rendah atau kadar toksisitas dengan tingkat kematian sel normal < 20% hanya terdapat pada senyawa observasi 1. Sedangkan 19 observasi lainnya memiliki kadar toksisitas yang

tinggi yang memiliki tingkat kematian pada sel normal antara 20% hingga 100 %. Berikut merupakan tabel tabulasi silang ANN *ensemble* dengan fungsi aktivasi *softmax.*

**Tabel 4.6** Tabulasi Silang pada *ANN Ensemble* Fungsi Aktivasi *Softmax*

| **Klasifikasi** | | | **Kelas Aktual** | | | |
|---|---|---|---|---|---|---|
| | | | ***Threshold* 0.3** | | ***Threshold* 0.5** | |
| | | | **Negatif** | **Positif** | **Negatif** | **Positif** |
| **Kelas Prediksi** | **12 *Subspace*** | **Negatif** | 18 | 1 | 18 | 2 |
| | | **Positif** | 0 | 1 | 0 | 0 |
| | **13 *Subspace*** | **Negatif** | 18 | 1 | 18 | 2 |
| | | **Positif** | 0 | 1 | 0 | 0 |
| | **14 *Subspace*** | **Negatif** | 18 | 1 | 18 | 2 |
| | | **Positif** | 0 | 1 | 0 | 0 |
| | **15 *Subspace*** | **Negatif** | 18 | 1 | 18 | 2 |
| | | **Positif** | 0 | 1 | 0 | 0 |
| | **16 *Subspace*** | **Negatif** | 18 | 2 | 18 | 2 |
| | | **Positif** | 0 | 0 | 0 | 0 |
| | **17 *Subspace*** | **Negatif** | 18 | 2 | 18 | 2 |
| | | **Positif** | 0 | 0 | 0 | 0 |
| | **18 *Subspace*** | **Negatif** | 18 | 2 | 18 | 2 |
| | | **Positif** | 0 | 0 | 0 | 0 |
| | **19 *Subspace*** | **Negatif** | 18 | 2 | 18 | 2 |
| | | **Positif** | 0 | 0 | 0 | 0 |

Pada tabel 4.6 terlihat bahwa tabulasi silang untuk 16 *subspace* hingga 19 *subspace* pada *threshold* optimal memiliki nilai yang sama dengan semua *subspace* pada *threshold* 0.5 yaitu dimana untuk terdapat 18 observasi yang tepat diprediksi ke dalam kadar toksisitas tinggi dan 0 observasi tepat diprediksi ke dalam kadar toksisitas rendah. Sedangkan pada *trheshold* optimum pada 12 hingga 19 *subspace* terdapat 18 observasi yang tepat diprediksi ke dalam kadar toksisitas tinggi dan 1 observasi yang tepat diprediksi ke dalam toksisitas rendah. Berikut merupakan nilai akurasi. *sensitivity*. *specificity*. dan AUC yang diperoleh.

**Tabel 4.7** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Softmax Threshold* Optimum

| Model | Accuracy | Sensitivity | Specificity | AUC |
|---|---|---|---|---|
| **12 Subspace** | **0.9500** | **0.5000** | **1.0000** | **0.7500** |
| **13 Subspace** | **0.9500** | **0.5000** | **1.0000** | **0.7500** |
| **14 *Subspace*** | **0.9500** | **0.5000** | **1.0000** | **0.7500** |
| **15 *Subspace*** | **0.9500** | **0.5000** | **1.0000** | **0.7500** |
| 16 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 17 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 18 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 19 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |

Angka bercetak tebal merupakan akurasi optimum

**Tabel 4.8** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Softmax Threshold* 0.5

| Model | *Accuracy* | *Sensitivity* | *Specificity* | AUC |
|---|---|---|---|---|
| 12 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 13 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 14 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 15 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 16 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 17 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 18 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 19 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |

Model *artificial neural network ensemble* dengan fungsi aktivasi *softmax* pada data *citotoxicity* dari 12 sampai 15 *subspace* dengan *threshold* optimum mampu memperoleh akurasi sebesar 95%. Dengan nilai *sensitivity* dan *specificity* sebesar 50% dan 100%. AUC yang dihasilkan oleh analisis ini adalah 75%. Sedangkan pada 16 sampai 19 *subspace* mampu memperoleh akurasi sebesar 90%. Dengan nilai *sensitivity* dan *specificity* sebesar 00% dan 100%. AUC yang dihasilkan oleh analisis ini adalah 50%.

Sedangkan pada *threshold* 0.5. didapatkan nilai akurasi sebesar 90% pada semua *subspace* . selain itu juga didapatkan nilai *sensitivity* dan *specificity* sebesar 0% dan 100% pada semua *subspace* . Pada *threshold* 0.5 nilai AUC diperoleh sebesar 50 %. Nilai *sensitifity* sebesar 0% menunjukkan bahwa dengan menggunakan *threshold* 0.5 maupun *threshold* optimum. metode yang digunakan tidak mampu memprediksi dengan tepat kelas *true postif* atau kadar *citotoxicity* yang rendah.

**b. *Sigmoid Activation Function***

Selain menggunakan fungsi aktivasi *softmax*. dalam penelitian ini juga menggunakan fungsi aktivasi *sigmoid* dimana pada fungsi aktivasi ini memiliki nilai pada *range* 0 sampai 1. Dengan menggunakan persamaan 4.5. threshold yang digunakan pada fungsi aktivasi sigmoid yaitu 0.3 dengan persamaan fungsi aktivasi sebagai berikut.

$$f(x) = \frac{1}{1+e^{-x}} \tag{4.10}$$

Dengan menggunakan threshold 0.3. maka didapatkan hasil probabilitas *artificial neural network ensemble* dengan fungsi aktivasi *sigmoid* yang dapat dilihat pada tabel berikut.

**Tabel 4.9** *Probability* Setiap *Subspace* untuk Fungsi Aktivasi *Sigmoid*

| **Obs** | **Probabilitas Tiap *Subspace*** | | | | **Kelas** |
|---|---|---|---|---|---|
| | **12** | **13** | **14** | **15** | |
| 1 | 0.2716(0) | 0.2725(0) | 0.2726(0) | 0.2727(0) | 0 |
| 2 | 0.1780(0) | 0.1652(0) | 0.1784(0) | 0.1640(0) | 0 |
| 3 | 0.1145(0) | 0.1153(0) | 0.1192(0) | 0.1234(0) | 0 |
| 4 | 0.1012(0) | 0.1024(0) | 0.1073(0) | 0.1146(0) | 0 |
| 5 | 0.0987(0) | 0.0995(0) | 0.1070(0) | 0.1074(0) | 0 |
| 6 | 0.1086(0) | 0.1065(0) | 0.1095(0) | 0.1193(0) | 0 |
| 7 | 0.1180(0) | 0.1177(0) | 0.1204(0) | 0.1190(0) | 0 |
| 8 | 0.0987(0) | 0.1020(0) | 0.1086(0) | 0.1128(0) | 0 |
| 9 | 0.0889(0) | 0.0973(0) | 0.1016(0) | 0.1064(0) | 0 |
| 10 | 0.1703(0) | 0.1769(0) | 0.1731(0) | 0.1649(0) | 0 |
| 11 | 0.0959(0) | 0.1021(0) | 0.1047(0) | 0.1080(0) | 0 |
| 12 | 0.1190(0) | 0.1186(0) | 0.1239(0) | 0.1202(0) | 0 |
| 13 | 0.0882(0) | 0.0882(0) | 0.0893(0) | 0.0968(0) | 0 |
| 14 | 0.2498(0) | 0.2243(0) | 0.2308(0) | 0.2191(0) | 0 |

**Tabel 4.9** *Probability* Setiap *Subspace* untuk Fungsi Aktivasi *Sigmoid* (lanjutan)

| Obs | Probabilitas Tiap *Subspace* | | | | Kelas |
|---|---|---|---|---|---|
| | **12** | **13** | **14** | **15** | |
| 15 | 0.1239(0) | 0.1209(0) | 0.1307(0) | 0.1337(0) | 0 |
| 16 | 0.0990(0) | 0.0980(0) | 0.1043(0) | 0.1020(0) | 0 |
| 17 | 0.1072(0) | 0.1072(0) | 0.1103(0) | 0.1173(0) | 0 |
| 18 | 0.0890(0) | 0.0901(0) | 0.0944(0) | 0.1023(0) | 0 |
| 19 | 0.0970(0) | 0.1041(0) | 0.1139(0) | 0.1155(0) | 0 |
| 20 | 0.1057(0) | 0.1120(0) | 0.1128(0) | 0.1200(0) | 0 |
| **Obs** | **Probabilitas Tiap *Subspace*** | | | | **Kelas** |
| | **16** | **17** | **18** | **19** | |
| 1 | 0.2728(0) | 0.2729(0) | 0.2730(0) | 0.2731(0) | 0 |
| 2 | 0.1641(0) | 0.1727(0) | 0.1627(0) | 0.1671(0) | 0 |
| 3 | 0.1227(0) | 0.1258(0) | 0.1320(0) | 0.1312(0) | 0 |
| 4 | 0.1147(0) | 0.1147(0) | 0.1233(0) | 0.1235(0) | 0 |
| 5 | 0.1086(0) | 0.1121(0) | 0.1123(0) | 0.1184(0) | 0 |
| 6 | 0.1253(0) | 0.1155(0) | 0.1233(0) | 0.1269(0) | 0 |
| 7 | 0.1263(0) | 0.1253(0) | 0.1349(0) | 0.1328(0) | 0 |
| 8 | 0.1206(0) | 0.1178(0) | 0.1220(0) | 0.1273(0) | 0 |
| 9 | 0.1157(0) | 0.1125(0) | 0.1178(0) | 0.1172(0) | 0 |
| 10 | 0.1746(0) | 0.1752(0) | 0.1696(0) | 0.1762(0) | 0 |
| 11 | 0.1098(0) | 0.1180(0) | 0.1115(0) | 0.1161(0) | 0 |
| 12 | 0.1315(0) | 0.1363(0) | 0.1264(0) | 0.1315(0) | 0 |
| 13 | 0.1034(0) | 0.1033(0) | 0.1078(0) | 0.1115(0) | 0 |
| 14 | 0.2044(0) | 0.2075(0) | 0.1992(0) | 0.2099(0) | 0 |
| 15 | 0.1357(0) | 0.1367(0) | 0.1443(0) | 0.1411(0) | 0 |
| 16 | 0.1050(0) | 0.1097(0) | 0.1145(0) | 0.1168(0) | 0 |
| 17 | 0.1207(0) | 0.1230(0) | 0.1228(0) | 0.1251(0) | 0 |
| 18 | 0.1038(0) | 0.1044(0) | 0.1079(0) | 0.1148(0) | 0 |
| 19 | 0.1222(0) | 0.1156(0) | 0.1174(0) | 0.1279(0) | 0 |
| 20 | 0.1170(0) | 0.1264(0) | 0.1287(0) | 0.1266(0) | 0 |

Pada tabel 4.9 terlihat bahwa untuk semua *subspace* dan observasi. tidak terdapat observasi yang masuk ke dalam kelas 1 yang terlihat dari prediksi pada tiap *subspace*. Dimana kelas 1 merupakan kelas dengan nilai probabilitas lebih dari 0.3. Sehingga. dengan menggunakan fungsi aktivasi sigmoid didapatkan semua observasi ke dalam kelas 0 atau ke dalam kelas kadar toksisitas tinggi (toksisitas dengan tingat kematian sel normal 20% - 100 %). Nilai pada tabel 4.9 diperoleh melalui persamaan berikut.

Sehingga persamaan ANN untuk fungsi *sigmoid* terbentuk sebagai berikut.

$$\hat{Y}(k) = \frac{1}{w_1^* x_1^* + w_2^* x_2^*} \tag{4.11}$$

Berikut merupakan tabulasi silang hasil klasifikasi *citotoxicity* dengan fungsi aktivasi *sigmoid*.

**Tabel 4.10** Tabulasi Silang pada Model ANN *Ensemble* Fungsi *Sigmoid*

| **Klasifikasi** | | | **Kelas Aktual** | | | |
|---|---|---|---|---|---|---|
| | | | ***Threshold* 0.3** | | ***Threshold* 0.5** | |
| | | | **Negatif** | **Positif** | **Negatif** | **Positif** |
| **Kelas Prediksi** | **12 *Subspace*** | **Negatif** | 18 | 2 | 18 | 2 |
| | | **Positif** | 0 | 0 | 0 | 0 |
| | **13 *Subspace*** | **Negatif** | 18 | 2 | 18 | 2 |
| | | **Positif** | 0 | 0 | 0 | 0 |
| | **14 *Subspace*** | **Negatif** | 18 | 2 | 18 | 2 |
| | | **Positif** | 0 | 0 | 0 | 0 |
| | **15 *Subspace*** | **Negatif** | 18 | 2 | 18 | 2 |
| | | **Positif** | 0 | 0 | 0 | 0 |
| | **16 *Subspace*** | **Negatif** | 18 | 2 | 18 | 2 |
| | | **Positif** | 0 | 0 | 0 | 0 |
| | **17 *Subspace*** | **Negatif** | 18 | 2 | 18 | 2 |
| | | **Positif** | 0 | 0 | 0 | 0 |
| | **18 *Subspace*** | **Negatif** | 18 | 2 | 18 | 2 |
| | | **Positif** | 0 | 0 | 0 | 0 |
| | **19 *Subspace*** | **Negatif** | 18 | 2 | 18 | 2 |
| | | **Positif** | 0 | 0 | 0 | 0 |

Melalui tabulasi silang pada tabel 4.10 terlihat bahwa nilai tabulasi untuk semua *subspace* pada kedua threshold memiliki nilai yang sama. Dimana terdapat 18 observasi yang tepat masuk ke dalam kadar toksisitas tinggi dan 0 obeservasi yang diprediksi tepat masuk ke dalam kadar toksisitas rendah atau tidak terdapat observasi dengan kadar toksisitas rendah. Sehingga tidak terdapat perbedaan pada kedua threshold. Maka melalui tabel berikut dilakukan perhitungan ketepatan klasifikasi.

**Tabel 4.11** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Sigmoid Threshold 0.3*

| **Model** | ***Accuracy*** | ***Sensitivity*** | ***Specificity*** | ***AUC*** |
|---|---|---|---|---|
| 12 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 13 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 14 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 15 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 16 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 17 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 18 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 19 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |

**Tabel 4.12** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Sigmoid Threshold 0.5*

| **Model** | **Accuracy** | **Sensitivity** | **Specificity** | **AUC** |
|---|---|---|---|---|
| 12 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 13 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 14 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 15 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 16 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 17 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 18 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |
| 19 *Subspace* | 0.9000 | 0.0000 | 1.0000 | 0.5000 |

Pada kedua *threshold*. baik 0.3 ataupun 0.5. dengan fungsi aktivasi *sigmoid* pada 12 sampai 19 *subspace* mampu memperoleh akurasi sebesar 90%. Dengan nilai *sensitivity* dan *specificity* sebesar 0% dan 100% untuk kedua *threshold*. Hal ini menunjukkan bahwa data *citotoxicity* dengan *neural network ensemble* menggunakan fungsi aktivasi tidak memiliki kemampuan untuk mengklasifikasikan *true positif* atau kadar *citotoxicity* yang rendah. Pada fungsi aktivasi *sigmoid* didapatkan AUC untuk kedua *threshold* dengan hasil yang sama yaitu sebesar 50%. Sehingga penggunaan fungsi aktivasi sigmoid dengan threshold optimum dan 0.5 memiliki ukuran kebaikan model yang sama

## B. ***Radioprotective***

### **a.** ***Softmax Activation Function***

Data *radioprotective* merupakan dimana data yang digunakan merupakan data dengan senyawa-senyawa gen p53 yang mampu memberikan radiasi optimum pada sel kanker. Dengan menggunakan persamaan fungsi aktivasi yang sama dengan data *citotoxicity* dan threshold optimum yang digunakan sesuai dengan persamaan 4.6. maka didapatkan nilai proabilitas *artificial neural network ensemble* dengan fungsi aktivasi *softmax* yang dapat dilihat pada tabel 4.13 berikut.

**Tabel 4.13** *Probability* Setiap *Subspace* untuk Fungsi Aktivasi *Softmax*

| **Obs** | **Probabilitas Tiap *Subspace*** | | | | **Kelas** |
|---|---|---|---|---|---|
| | **12** | **13** | **14** | **15** | |
| 1 | 0.5600249(0) | 0.563144(0) | 0.5633433(0) | 0.5781855(0) | 0 |
| 2 | 0.7603796(1) | 0.7548905(1) | 0.7463678(1) | 0.7396479(1) | 1 |
| 3 | 0.7849111(1) | 0.7833242(1) | 0.7820656(1) | 0.76493(1) | 1 |
| 4 | 0.7500233(1) | 0.7521468(1) | 0.7469708(1) | 0.7457432(1) | 1 |
| 5 | 0.7840135(1) | 0.7742418(1) | 0.7657103(1) | 0.7643068(1) | 1 |
| 6 | 0.7704249(1) | 0.7628475(1) | 0.7744716(1) | 0.7758982(1) | 1 |
| 7 | 0.7381216(1) | 0.7420703(1) | 0.7453186(1) | 0.7420778(1) | 1 |
| 8 | 0.8130632(1) | 0.8183108(1) | 0.8121869(1) | 0.8097874(1) | 1 |
| 9 | 0.8162464(1) | 0.8142361(1) | 0.8140293(1) | 0.8091604(1) | 1 |
| 10 | 0.5357344(0) | 0.5341888(0) | 0.5489356(0) | 0.5625659(0) | 0 |
| 11 | 0.7843958(1) | 0.7791211(1) | 0.7790264(1) | 0.7801213(1) | 1 |
| 12 | 0.645071(1) | 0.6698062(1) | 0.6563395(1) | 0.6557025(1) | 1 |
| 13 | 0.6985157(1) | 0.7089837(1) | 0.7042385(1) | 0.7175963(1) | 1 |
| 14 | 0.7491571(1) | 0.7250287(1) | 0.7341551(1) | 0.7394824(1) | 1 |
| 15 | 0.8172933(1) | 0.7915302(1) | 0.8190695(1) | 0.8024668(1) | 1 |
| 16 | 0.6860529(1) | 0.6971996(1) | 0.6922063(1) | 0.7067655(1) | 1 |
| 17 | 0.8355765(1) | 0.8288243(1) | 0.8258853(1) | 0.8199382(1) | 1 |
| 18 | 0.8049749(1) | 0.7967266(1) | 0.7842352(1) | 0.7868724(1) | 1 |
| 19 | 0.7726505(1) | 0.7720608(1) | 0.7668453(1) | 0.7747456(1) | 1 |
| 20 | 0.8025978(1) | 0.8050988(1) | 0.7885905(1) | 0.7971036(1) | 1 |

| **Obs** | **Probabilitas Tiap *Subspace*** | | | | **Kelas** |
|---|---|---|---|---|---|
| | **16** | **17** | **18** | **19** | |
| 1 | 0.601148(0) | 0.5989991(0) | 0.6089979(0) | 0.633318(1) | 0 |
| 2 | 0.7282242(1) | 0.7417105(1) | 0.7478415(1) | 0.7400526(1) | 1 |

**Tabel 4.13** *Probability* Setiap *Subspace* untuk Fungsi Aktivasi *Softmax* *(lanjutan)*

| Obs | Probabilitas Tiap *Subspace* | | | | Kelas |
|---|---|---|---|---|---|
| | 16 | 17 | 18 | 19 | |
| 3 | 0.7748862(1) | 0.7668998(1) | 0.7610379(1) | 0.7644367(1) | 1 |
| 4 | 0.7494252(1) | 0.7405827(1) | 0.746929(1) | 0.7407779(1) | 1 |
| 5 | 0.7626621(1) | 0.7709203(1) | 0.7638829(1) | 0.7716454(1) | 1 |
| 6 | 0.7628202(1) | 0.7594645(1) | 0.7549914(1) | 0.754876(1) | 1 |
| 7 | 0.7388775(1) | 0.7454211(1) | 0.7462164(1) | 0.7431474(1) | 1 |
| 8 | 0.8055043(1) | 0.790701(1) | 0.7976706(1) | 0.794237(1) | 1 |
| 9 | 0.8073454(1) | 0.7949374(1) | 0.798719(1) | 0.7943076(1) | 1 |
| 10 | 0.5722803(0) | 0.5912838(0) | 0.6034536(0) | 0.6198981(0) | 0 |
| 11 | 0.7649411(1) | 0.7776958(1) | 0.7705085(1) | 0.7773666(1) | 1 |
| 12 | 0.6600973(1) | 0.678215(1) | 0.6855149(1) | 0.7015186(1) | 1 |
| 13 | 0.716587(1) | 0.70844(1) | 0.7122233(1) | 0.7161785(1) | 1 |
| 14 | 0.728304(1) | 0.7363372(1) | 0.7407987(1) | 0.7339036(1) | 1 |
| 15 | 0.8013912(1) | 0.7881975(1) | 0.7806295(1) | 0.7902681(1) | 1 |
| 16 | 0.7100673(1) | 0.7066638(1) | 0.7207685(1) | 0.7158924(1) | 1 |
| 17 | 0.8173347(1) | 0.8071362(1) | 0.8056858(1) | 0.7993897(1) | 1 |
| 18 | 0.7869178(1) | 0.7821306(1) | 0.7792272(1) | 0.7827023(1) | 1 |
| 19 | 0.7677083(1) | 0.7568864(1) | 0.7680791(1) | 0.7661518(1) | 1 |
| 20 | 0.7946544(1) | 0.7787746(1) | 0.7852435(1) | 0.7734156(1) | 1 |

Berdasarkan hasil prediksi pada tabel 4.13. diketahui bahwa terdapat 2 observasi yang diduga memiliki proteksi radiasi yang rendah dimana tingkat kematian sel kanker ≤10 %. Selain itu. juga diketahui bahwa sejumlah 18 observasi lainnya diduga memiliki proteksi radiasi yang tinggi dengan kematian sel kanker >10%. Nilai-nilai probabilitas di atas didapatkan melalui persamaan berikut.

$$f(x_i) = \frac{\exp(x_i)}{\sum_{k=1}^{p} \exp(x_k)} \tag{4.8}$$

Dengan p merupakan banyak jumlah node pada setiap *subspace*. Sehingga persamaan ANN untuk fungsi *softmax* terbentuk sebagai berikut.

$$\hat{Y}_1(k) = \frac{w_1^* x_1^* + w_3^* x_2^*}{w_1^* x_1^* + w_2^* x_2^* + w_3^* x_3^* + w_4^* x_4^*} \qquad (4.9)$$

$$\hat{Y}_2(k) = \frac{w_2^* x_1^* + w_4^* x_2^*}{w_1^* x_1^* + w_2^* x_2^* + w_3^* x_3^* + w_4^* x_4^*} \qquad (4.10)$$

Dimana,
w* = weight pada output layer
x* = node input yang telah diolah pada hidden layer

Untuk memastikan lebih lanjut. maka dilakukan penghitungan melalui tabulasi silang. Berikut merupakan tabulasi silang hasil klasifikasi.

**Tabel 4.14** Tabulasi Silang pada Model ANN *Ensemble* Fungsi *Softmax*

| **Klasifikasi** | | | **Kelas Aktual** | | | |
|---|---|---|---|---|---|---|
| | | | ***Threshold* 0.625** | | ***Threshold* 0.5** | |
| | | | **Negatif** | **Positif** | **Negatif** | **Positif** |
| **Kelas prediksi** | **12 *Subspace*** | **Negatif** | 2 | 0 | 0 | 0 |
| | | **Positif** | 3 | 15 | 5 | 15 |
| | **13 *Subspace*** | **Negatif** | 2 | 0 | 0 | 0 |
| | | **Positif** | 3 | 15 | 5 | 15 |
| | **14 *Subspace*** | **Negatif** | 2 | 0 | 0 | 0 |
| | | **Positif** | 3 | 15 | 5 | 15 |
| | **15 *Subspace*** | **Negatif** | 2 | 0 | 0 | 0 |
| | | **Positif** | 3 | 15 | 5 | 15 |
| **Kelas prediksi** | **16 *Subspace*** | **Negatif** | 2 | 0 | 0 | 0 |
| | | **Positif** | 3 | 15 | 5 | 15 |
| | **17 *Subspace*** | **Negatif** | 2 | 0 | 0 | 0 |
| | | **Positif** | 3 | 15 | 5 | 15 |
| | **18 *Subspace*** | **Negatif** | 2 | 0 | 0 | 0 |
| | | **Positif** | 3 | 15 | 5 | 15 |
| | **19 *Subspace*** | **Negatif** | 1 | 0 | 0 | 0 |
| | | **Positif** | 4 | 15 | 5 | 15 |

Dari hasil tabel 4.14 diketahui bahwa untuk threshold optimum memiliki nilai tabulasi yang sama untuk *subspace* 12 hingga 18. yaitu terdapat 2 observasi yang tepat diprediksi ke dalam kelas *radioprotective* rendah dan sebanyak 15 observasi yang tepat diprediksi ke dalam kelas *radioprotective* tinggi. Sedangkan untuk 19 *subspace,* nilai tabulasinya menunjukkan bahwa terdapat 1 observasi yang tepat diprediksi ke dalam kelas *radioprotective* rendah dan 15 observasi yang tepat diprediksi ke dalam kelas *radioprotective* tinggi.

Sedangkan pada *threshold* 0.5. diketahui bahwa tidak terdapat observasi yang tepat diprediksi ke dalam kelas *radioprotective* rendah dan sejumlah 15 observasi tepat diprediksi ke dalam kelas *radioprotective* tinggi. Untuk melihat ketepatan hasil klasifikasi. maka dapat dilihat pada tabel berikut.

**Tabel 4.15** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Softmax Threshold 0.625*

| Model | *Accuracy* | *Sensitivity* | *Specificity* | *AUC* |
|---|---|---|---|---|
| **12 *Subspace*** | **0.8500** | **1.0000** | **0.4000** | **0.7000** |
| **13 *Subspace*** | **0.8500** | **1.0000** | **0.4000** | **0.7000** |
| **14 *Subspace*** | **0.8500** | **1.0000** | **0.4000** | **0.7000** |
| **15 *Subspace*** | **0.8500** | **1.0000** | **0.4000** | **0.7000** |
| **16 *Subspace*** | **0.8500** | **1.0000** | **0.4000** | **0.7000** |
| **17 *Subspace*** | **0.8500** | **1.0000** | **0.4000** | **0.7000** |
| **18 *Subspace*** | **0.8500** | **1.0000** | **0.4000** | **0.7000** |
| 19 *Subspace* | 0.8000 | 1.0000 | 0.2000 | 0.6000 |

*Angka bercetak tebal merupakan akurasi optimum*

**Tabel 4.16** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Softmax Threshold 0.5*

| Model | *Accuracy* | *Sensitivity* | *Specificity* | *AUC* |
|---|---|---|---|---|
| 12 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 13 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 14 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 15 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 16 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |

**Tabel 4.16** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Softmax Threshold 0.5 (lanjutan)*

| Model | Accuracy | Sensitivity | Specificity | AUC |
|---|---|---|---|---|
| 17 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 18 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 19 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |

Model *ANN ensemble* dengan fungsi aktivasi *softmax* dengan *threshold* 0.625 pada 12 sampai 18 *subspace* mampu memperoleh akurasi sebesar 85%. Dengan nilai *sensitivity* dan *specificity* sebesar 100% dan 40%. AUC yang dihasilkan oleh fungsi aktivasi ini adalah 70%. Sedangkan pada 19 *subspace* mampu memperoleh akurasi sebesar 80%. Pada *threshold* 0.5 didapatkan nilai *sensitivity* dan *specificity* sebesar 100% dan 0%. Nilai sensitifity sebesar 0% menunjukkan bahwa fungsi aktivasi *softmax* dengan *threshold* 0.5 tidak mampu memprediksi dengan tepat kelas *true negatif* atau kadar *radioprotective* yang rendah.

AUC yang dihasilkan oleh analisis ini adalah 50%. Nilai akurasi data *radioprotective* dengan fungsi aktivasi *softmax* menggunakan *threshold* 0.5 memiliki nilai lebih kecil daripada menggunakan *threshold* optimum. yaitu sebesar 75%.

**b. *Sigmoid Activation Function***

Dengan menggunakan *threshold* optimum sesuai dengan persamaan 4.6. maka didapatkan nilai probabilitas pada tiap *subspace* pada data *radioprotective* dengan menggunakan fungsi aktivasi sigmoid sebagai berikut.

**Tabel 4.17** *Probability* Setiap *Subspace* untuk Fungsi Aktivasi *Sigmoid*

| **Obs** | **Probabilitas Tiap *Subspace*** | | | | **Kelas** |
|---|---|---|---|---|---|
| | **12** | **13** | **14** | **15** | |
| 1 | 0.6073(0) | 0.5631(0) | 0.6174 (0) | 0.6106 (0) | 0 |
| 2 | 0.7393 (1) | 0.75489 (1) | 0.7324 (1) | 0.7364 (1) | 1 |
| 3 | 0.7611(1) | 0.7833 (1) | 0.7495 (1) | 0.7610 (1) | 1 |
| 4 | 0.7458(1) | 0.7521 (1) | 0.7450 (1) | 0.7409 (1) | 1 |
| 5 | 0.7627(1) | 0.7742 (1) | 0.7621 (1) | 0.7627 (1) | 1 |
| 6 | 0.7579(1) | 0.7628 (1) | 0.7565 (1) | 0.7516 (1) | 1 |
| 7 | 0.7365(1) | 0.7420 (1) | 0.7425(1) | 0.7380 (1) | 1 |

**Tabel 4.17** *Probability* Setiap *Subspace* untuk Fungsi Aktivasi *Sigmoid* (lanjutan)

| | | | | | |
|---|---|---|---|---|---|
| 8 | 0.7938(1) | 0.8183 (1) | 0.7845 (1) | 0.7859 (1) | 1 |
| 9 | 0.7981(1) | 0.8142 (1) | 0.7925 (1) | 0.7867 (1) | 1 |
| 10 | 0.5902(0) | 0.5341 (0) | 0.6221 (0) | 0.6100 (0) | 0 |
| 11 | 0.7660(1) | 0.7791 (1) | 0.7637 (1) | 0.7686 (1) | 1 |
| 12 | 0.6863 (1) | 0.6698 (1) | 0.6888 (1) | 0.6866 (1) | 1 |
| 13 | 0.7227 (1) | 0.7089 (1) | 0.7183 (1) | 0.7126 (1) | 1 |
| 14 | 0.7383 (1) | 0.7250 (1) | 0.7238 (1) | 0.7250 (1) | 1 |
| 15 | 0.7869 (1) | 0.7915 (1) | 0.7691 (1) | 0.7783 (1) | 1 |
| 16 | 0.7096(1) | 0.6971 (1) | 0.7057 (1) | 0.7027 (1) | 1 |
| 17 | 0.8073 (1) | 0.8288 (1) | 0.7872 (1) | 0.7926 (1) | 1 |
| 18 | 0.7839 (1) | 0.7967 (1) | 0.7848 (1) | 0.7738 (1) | 1 |
| 19 | 0.7789 (1) | 0.7720 (1) | 0.7626 (1) | 0.7536 (1) | 1 |
| 20 | 0.7852 (1) | 0.8050 (1) | 0.7686 (1) | 0.7794 (1) | 1 |

| **Obs** | **Probabilitas Tiap *Subspace*** | | | | **Kelas** |
|---|---|---|---|---|---|
| | **16** | **17** | **18** | **19** | |
| 1 | 0.6308 (1) | 0.6453(1) | 0.649(1) | 0.662 (1) | 0 |
| 2 | 0.7323 (1) | 0.7262 (1) | 0.733(1) | 0.727 (1) | 1 |
| 3 | 0.7603 (1) | 0.7500(1) | 0.753(1) | 0.749 (1) | 1 |
| 4 | 0.7368 (1) | 0.7333(1) | 0.739 (1) | 0.733 (1) | 1 |
| 5 | 0.7599 (1) | 0.7554(1) | 0.754(1) | 0.747 (1) | 1 |
| 6 | 0.7527 (1) | 0.7527(1) | 0.7518 (1) | 0.743 (1) | 1 |
| 7 | 0.7367 (1) | 0.7310(1) | 0.735 (1) | 0.725 (1) | 1 |
| 8 | 0.7777 (1) | 0.7733(1) | 0.771 (1) | 0.767 (1) | 1 |
| 9 | 0.7798 (1) | 0.7749(1) | 0.772 (1) | 0.763 (1) | 1 |
| 10 | 0.6289 (1) | 0.6374(1) | 0.639 (1) | 0.647 (1) | 0 |
| 11 | 0.7667 (1) | 0.7638(1) | 0.760 (1) | 0.755 (1) | 1 |
| 12 | 0.6991 (1) | 0.6950(1) | 0.695 (1) | 0.704 (1) | 1 |
| 13 | 0.7232 (1) | 0.7264(1) | 0.721 (1) | 0.725 (1) | 1 |
| 14 | 0.7350 (1) | 0.7221(1) | 0.7242 (1) | 0.730 (1) | 1 |
| 15 | 0.7683 (1) | 0.7595(1) | 0.757 (1) | 0.754 (1) | 1 |
| 16 | 0.7233 (1) | 0.7231(1) | 0.727 (1) | 0.715 (1) | 1 |
| 17 | 0.7782 (1) | 0.7762(1) | 0.772 (1) | 0.767 (1) | 1 |
| 18 | 0.7688 (1) | 0.7679(1) | 0.766 (1) | 0.756 (1) | 1 |
| 19 | 0.7572 (1) | 0.7599(1) | 0.752(1) | 0.747 (1) | 1 |
| 20 | 0.7618 (1) | 0.7726(1) | 0.7575 (1) | 0.759 (1) | 1 |

Berdasarkan nilai probabilitas pada tabel 4.17. diketahui bahwa hanya terdapat 2 observasi yang masuk ke dalam kelas 0 atau kelas *radioprotective* rendah. Dengan menggunakan fungsi aktivasi sigmoid pada persamaan 4.10, maka didapatkan nilai probabilitas dengan persamaan 4.11 berikut.

$$f(x) = \frac{1}{1+e^{-x}} \tag{4.10}$$

$$\hat{Y}(k) = \frac{1}{w_1^* x_1^* + w_2^* x_2^*} \tag{4.11}$$

Untuk memastikan hasil klasifikasi tersebut maka dilakukan analisis berdasarkan tabulasi silangnya. Berikut ini merupakan tabulasi silang hasil klasifikasi.

**Tabel 4.18** Tabulasi Silang pada Model ANN *Ensemble* Fungsi *Sigmoid*

| **Klasifikasi** | | | **Kelas Aktual** | | | |
|---|---|---|---|---|---|---|
| | | | ***Threshold* 0.625** | | ***Threshold* 0.5** | |
| | | | **Negatif** | **Positif** | **Negatif** | **Positif** |
| **Kelas prediksi** | **12 *Subspace*** | **Negatif** | 2 | 0 | 0 | 0 |
| | | **Positif** | 3 | 15 | 5 | 15 |
| | **13 *Subspace*** | **Negatif** | 2 | 0 | 0 | 0 |
| | | **Positif** | 3 | 15 | 5 | 15 |
| | **14 *Subspace*** | **Negatif** | 2 | 0 | 0 | 0 |
| | | **Positif** | 3 | 15 | 5 | 15 |
| | **15 *Subspace*** | **Negatif** | 2 | 0 | 0 | 0 |
| | | **Positif** | 3 | 15 | 5 | 15 |
| | **16 *Subspace*** | **Negatif** | 0 | 0 | 0 | 0 |
| | | **Positif** | 5 | 15 | 5 | 15 |
| | **17 *Subspace*** | **Negatif** | 0 | 0 | 0 | 0 |
| | | **Positif** | 5 | 15 | 5 | 15 |
| | **18 *Subspace*** | **Negatif** | 0 | 0 | 0 | 0 |
| | | **Positif** | 5 | 15 | 5 | 15 |
| | **19 *Subspace*** | **Negatif** | 0 | 0 | 0 | 0 |
| | | **Positif** | 5 | 15 | 5 | 15 |

Dari tabulsi silang di atas. didapatkan hasil bahwa untuk threshold 0.625 pada 12 hingga 15 *subspace* diketahui terdapat 2 observasi yang diprediksi tepat masuk ke dalam kelas radio proteksi rendah dan sebanyak 15 observasi diprediksi tepat masuk ke dalam kelas radio proteksi tinggi. Berikut merupakan perhitungan ketepatan klasifikasi.

**Tabel 4.19** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Sigmoid Threshold 0.625*

| Model | *Accuracy* | *Sensitivity* | *Specificity* | *AUC* |
|---|---|---|---|---|
| **12 *Subspace*** | **0.8500** | **1.0000** | **0.4000** | **0.7000** |
| **13 *Subspace*** | **0.8500** | **1.0000** | **0.4000** | **0.7000** |
| **14 *Subspace*** | **0.8500** | **1.0000** | **0.4000** | **0.7000** |
| **15 *Subspace*** | **0.8500** | **1.0000** | **0.4000** | **0.7000** |
| 16 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 17 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 18 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 19 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |

Angka bercetak tebal merupakan akurasi optimum

**Tabel 4.20** Ukuran Kebaikan Model ANN *Ensemble* Fungsi *Sigmoid Threshold 0.5*

| Model | *Accuracy* | *Sensitivity* | *Specificity* | *AUC* |
|---|---|---|---|---|
| 12 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 13 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 14 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 15 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 16 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 17 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 18 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |
| 19 *Subspace* | 0.7500 | 1.0000 | 0.0000 | 0.5000 |

Dari tabel di atas diketahui bahwa dengan fungsi aktivasi sigmoid pada *threshold* 0.625 untuk 12 sampai 15 *subspace* mampu memperoleh akurasi sebesar 85 % sedangkan pada

*threshold* 0.5 sebesar 75%. Dengan nilai *sensitivity* dan *specificity* pada *threshold* 0.625 sebesar 100 % dan 40 % sedangkan pada *threshold* 0.5 sebesar 100% dan 0%. Dengan menggunakan threshold 0.5 menunjukkan bahwa metode tidak mampu memrediksi dengan tepat kadar radioprotective rendah. Hal tersebut ditunjukan oleh nilai spesificity yang bernilai 0%.

Nilai AUC yang dihasilkan oleh kedua *threshold* pun berbeda. didapatkan nilai AUC sebesar 70 % pada *threshold* optimum untuk 12 hingga 15 *subspace* . Sedangkan pada 16 hingga 19 *subspace* hanya mampu memperoleh akurasi sebesar 50%. Dan pada *threshold* 0.5. didapatkan nilai AUC sebesar 50%.

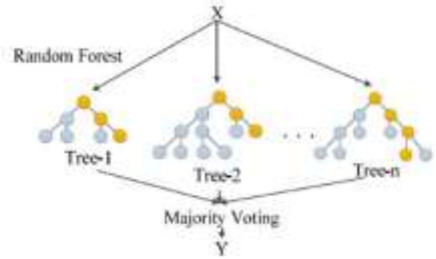
## 4.3 Klasifikasi Menggunakan *Random forest*

Untuk memudahkan pembacaan alur proses *random forest,* maka diberikan ilustrasi proses metode *random forest* sebagai berikut

### 4.3.1 Ilustrasi *Random Forest*

Bagian ini menjelaskan bagaimana *random forest* melakukan klasifikasi terhadap data *citotoxicity*. Berikut diberikan illustrasi mengenai langkah melakukan klasifikasi menggunakan metode *random forest* untuk data *citotoxicity*.

1. Melakukan penarikan data secara acak dari data *citotoxicity*.
2. Membentuk pohon hingga mencapai ukuran maksimum. Pembangunan pohon dilakukan dengan menerapkan random *feature selection* pada setiap proses pemilihan pemilah, yaitu $m$ variabel prediktor dipilih secara acak dimana $m << p$, lalu pemilah terbaik dipilih berdasarkan $m$ variabel prediktor tersebut.
3. Kemudian mengulangi langkah 1 dan 2 hingga terbentuk sebuah hutan. Dari beberapa pohon yang membentuk hutan tersebut didapatkan kelas prediksi melalui probabilitas yang dihasilkan, lalu dilakukan *majority voting* untuk menentukan kelas klasifikasinya. Bila digambarkan akan berbentuk sebagai berikut




**Gambar 4.9** Ilustrasi *Random Forest*

Setelah diketahui kelas terbanyak pada *majority voting* maka semua probabilitas pada kelas terbanyak hasil *majority voting* tersebut dihitung rata-ratanya.

4. Setelah didapat sebuah hutan dan probabilitas dari masingmasing pohon, kemudian menentukan *treshold* untuk menentukan kelas probabilitas. Untuk random forest didapatkan nilai probabilitas sebagai berikut.

$$threshold = \frac{p+0,5}{2} = \frac{0,1+0,5}{2} = 0,3 \qquad (4.12)$$

Sehingga masing-masing kelas observasi terlihat pada tabel berikut

**Tabel 4.21** *Probability* dan Kelas Prediksi Data *Citotoxicity*

| obs | ***Citotoxicity*** | |
|---|---|---|
| | ***Probability*** | **Kelas Prediksi** |
| 1 | 0,1512 | 0 |
| 2 | 0,3443 | 1 |
| 3 | 0,0437 | 0 |
| 4 | 0,0471 | 0 |
| 5 | 0,0233 | 0 |
| 6 | 0,0412 | 0 |
| 7 | 0,0652 | 0 |
| 8 | 0,0000 | 0 |
| 9 | 0,0057 | 0 |

**Tabel 4.21** *Probability* dan Kelas Prediksi Data *Citotoxicity (lanjutan)*

| **obs** | ***Citotoxicity*** | |
|---|---|---|
| | ***Probability*** | **Kelas Prediksi** |
| 10 | 0,5990 | 1 |
| 11 | 0,0282 | 0 |
| 12 | 0,1445 | 0 |
| 13 | 0,0452 | 0 |
| 14 | 0,2022 | 0 |
| 15 | 0,0281 | 0 |
| 16 | 0,0426 | 0 |
| 17 | 0,0201 | 0 |
| 18 | 0,0162 | 0 |
| 19 | 0,0170 | 0 |
| 20 | 0,0449 | 0 |

### 4.3.2 Klasifikasi Senyawa Penyusun Obat Kanker Menggunakan *Random Forest*

Analisis menggunakan metode ini berguna untuk menggambarkan hubungan antara variabel respon *cytotoxycity* dan *radioprotective*dengan 217 variabel prediktornya. Evaluasi dalam analisis ini mmenggunakan prosedur *Full Training Set*. Dimana keseluruhan data akan digunakan sebagai data *training* dan keseluruhan data pula yang digunakan sebagai data *testing* untuk menguji seberapa baik model yang terbentuk. Model *random forest* dibentuk dengan 500 *trees* dan variabel prediktor dipecah (*split*) sejumlah 72. *Threshold* yang digunakan dalam klasifikasi dengan respon biner adalah 0.3. Berikut merupakan rumus untuk menghitung *threshold* optimal untuk data *citotoxicity* dan *radioprotective*.

$$threshold\ citotoxicity = \frac{p+0,5}{2} = \frac{0,1+0,5}{2} = 0,3 \qquad (4.12)$$

$$threshold\ radioprotective = \frac{p+0,5}{2} = \frac{0,75+0,5}{2} = 0,625 \tag{4.13}$$

Dimana *p* adalah probabilitas pengamatan yang berada di kelas positif. Berikut ini hasil prediksi *probability* dan prediksi kelas masing-masing observasi.

**Tabel 4.22** *Probability* dan Kelas Prediksi

| obs | ***Citotoxicity*** | | ***Radioprotective*** | |
|---|---|---|---|---|
| | ***Probability*** | **Kelas Prediksi** | ***Probability*** | **Kelas Prediksi** |
| 1 | 0,1512 | 0 | 0,6448307 | 1 |
| 2 | 0,3443 | 1 | 0,7458 | 1 |
| 3 | 0,0437 | 0 | 0,8542 | 1 |
| 4 | 0,0471 | 0 | 0,6722 | 1 |
| 5 | 0,0233 | 0 | 0,5864 | 0 |
| 6 | 0,0412 | 0 | 0,6840 | 1 |
| 7 | 0,0652 | 0 | 0,6571 | 1 |
| 8 | 0,0000 | 0 | 0,8998 | 1 |
| 9 | 0,0057 | 0 | 0,8996 | 1 |
| 10 | 0,5990 | 1 | 0,6221 | 0 |
| 11 | 0,0282 | 0 | 0,8102 | 1 |
| 12 | 0,1445 | 0 | 0,8027 | 1 |
| 13 | 0,0452 | 0 | 0,8982 | 1 |
| 14 | 0,2022 | 0 | 0,3015 | 0 |
| 15 | 0,0281 | 0 | 0,7546 | 1 |
| 16 | 0,0426 | 0 | 0,9637 | 1 |
| 17 | 0,0201 | 0 | 0,8933 | 1 |
| 18 | 0,0162 | 0 | 0,8351 | 1 |
| 19 | 0,0170 | 0 | 0,8050 | 1 |
| 20 | 0,0449 | 0 | 0,6441 | 1 |

Hasil klasifikasi dengan menggunakan *random forest* pada tabel 4.21 terlihat bahwa pada *citotoxicity* terdapat 2 observasi yang masuk ke dalam kelas *citotoxicity* rendah dimana tingkat kematian sel normal kurang dari 20%. Selainnya. 18 observasi masuk ke dalam kelas *citotoxicity* tinggi dimana tingkat kematian sel normal berada pada kisaran persentasi 20% - 100%. Sedangkan untuk *radioprotective.* diketahui bahwa terdapat 18 observasi masuk ke dalam kelas 1 yang berarti bahwa tingkat proteksi radiasi nya tinggi dimana tingat kematian sel kanker kurang dari 10 %. Berikut ini merupakan tabulasi silang hasil klasifikasi.

**Tabel 4.23** Tabulasi Silang pada *Random forest*

| **Klasifikasi** | | **Kelas Aktual** | | | |
|---|---|---|---|---|---|
| | | ***Citotoxicity*** | | ***Radioprotective*** | |
| | | **Negatif** | **Positif** | **Negatif** | **Positif** |
| **Kelas Aktual** | **Negatif** | 16 | 2 | 1 | 2 |
| | **Positif** | 2 | 0 | 4 | 13 |

Berdasarkan tabel 4.22 dapat diketahui bahwa dari 20 data yang digunakan. terdapat 16 observasi tepat terprediksi kedalam kelas *citotoxicity* tinggi (normal 20 - 100%). 0 observasi tepat terprediksi kedalam kelas *citotoxicity* rendah (tingat kematian sel normal < 20 %) dan sisanya tidak terklasifikasi dengan tepat. Sedangkan untuk *radioprotection*. terdapat 1 observasi yang tepat terprediksi dalam kelas *radioprotective* rendah ( kematian sel kanker ≤ 10%). Terdapat 13 observasi tepat terprediksi dalam kelas *radioprotective* tinggi. dan sisanya tidak dapat terklasifikasi dengan tepat. Dari tabulasi diatas. selanjutnya dapat dhitung nilai keteatan klasifikasi. Berikut ini merupakan perhitungan ketepatan klasifikasi.

**Tabel 4.24** Ukuran Kebaikan Model *Random forest*

| **Model** | ***Accuracy*** | ***Sensitivity*** | ***Specificity*** | ***AUC*** |
|---|---|---|---|---|
| ***Citotoxicity*** | 0,8000 | 0,0000 | 0,8889 | 0,4445 |
| ***Radioprotective*** | 0,7000 | 0,8667 | 0,2000 | 0,5334 |

Model *Random forest* pada data *citotoxicity* mampu memperoleh akurasi sebesar 80%. sedangkan pada data *radioprotective* sebesar 70%. Dan didapatkan nilai *sensitivity* dan *specificity* sebesar 0% dan 88,89% pada *citotoxicity* dan sebesar 86,67% dan 20% pada *radioprotection*. Nilai *sensitifity* pada *citotoxicity* yang bernilai 0% menunjukkan bahwa metode *random forest* tidak mampu memprediksikan dengan tepat kelas *true positif* atau kadar *citotoxicity* yang rendah. Selain itu. nilai AUC yang dihasilkan pada data *citotoxicity* lebih rendah bila dibandingkan dengan data *radioprotective*yaitu sebesar 44,45%.

# BAB V
# KESIMPULAN DAN SARAN

## 5.1 Kesimpulan

Berdasarkan analisis dan pembahasan yang telah dilakukan pada bab 4. maka diperoleh kesimpulan sebagai berikut.

1. Klasifikasi senyawa terpenting pada *citotoxicity* dan *radioprotective* dengan menggunakan *neural network ensemble* menunjukkan bahwa terdapat beberapa senyawa observasi yang berbeda dengan kelas aktualnya. Pada data *citotoxicity* dengan menggunakan fungsi aktivasi softmax menunjukkan terdapat 1 observasi prediksi yang berbeda dengan kelas aktualnya. sedangkan untuk fungsi aktivasi sigmoid terdapat 2 observasi prediksi yang berbeda. Pada data *radioprotective* dengan fungsi aktivasi *softmax* dan *sigmoid* didapatkan 5 observasi prediksi yang berbeda. Untuk klasifikasi senyawa dengan menggunakan *random forest* didapatkan hasil bahwa terdapat 4 observasi prediksi yang berbeda pada data *citototxicity* dan 2 observasi prediksi yang sama dengan kelas aktualnya apada data *radioprotective*.
2. Hasil ketepatan klasifikasi dengan menggunakan *neural network ensemble* pada fungsi aktivasi *softmax* dan *sigmoid* memiliki nilai akurasi yang lebih tinggi pada fungsi aktivasi *softmax*, yaitu sebesar 95% pada data *citotoxicity*. Nilai akurasi yang tinggi ini didukung dengan nilai *sensitivity* dan *spesificity* sebesar 50% dan 100% serta nilai AUC sebesar 75%. Pada data *radioprotective* menunjukkan bahwa hasil ketepatan klasifikasi dengan kedua fungsi aktivasi memiliki nilai yang sama, yaitu dengan nilai akurasi sebesar 85% dan nilai *sensitivity* serta *spesificity* sebesar 100% dan 40% juga nilai AUC sebesar 70%. Sedangkan pada *random forest*. didapatkan nilai akurasi sebesar 80% untuk data *citototxicity* dan 70% untuk data *radioprotective*. Nilai *sensitivity* dan *spesificity* untuk *random forest* didapat sebesar 0% dan

88.89% pada *citotoxicity*. sedangkan untuk *radioprotective* didapatkan sebesar 86.67% dan 20%. Pada kedua data dengan menggunakan *random forest* didapatkan nilai AUC sebesar 50% dan 50%. Sehingga didapatkan metode terbaik untuk evaluasi klasifikasi senyawa penyusun obat kanker adalah menggunakan *artificial neural network ensemble* dengan fungsi aktivasi *softmax*.

3. Pada *neural network ensemble* pada kedua fungsi aktivasi, *softmax* dan *sigmoid*, diketahui bahwa semakin besar jumlah *subspace* maka nilai ketepatan klasifikasi diduga kurang baik. Pada penelitian terlihat bahwa nilai ketepatan klasifikasi untuk 12 hingga 15 *subspace* memiliki nilai lebih tinggi bila dibandingkan dengan *subspace* selanjutnya.

## 5.2 Saran

Berdasarkan kesimpulan yang diperoleh, didapatkan beberapa saran sebagai pertimbangan penelitian selanjutnya adalah sebagai berikut.

1. Pada metode *artificial neural network ensemble* disarankan bagi penelitian selanjutnya untuk mencoba jumlah *node* pada *hidden layer* dari jumlah node terkecil hingga *node* optimumnya.
2. Dilakukan pengkajian ulang terkait metode yang digunakan agar didapatkan hasil yang memiliki ketepatan klasifikasi lebih tinggi.

# DAFTAR PUSTAKA

Agrawal, S., & Agrawal, J. (2015). Neural Network Techniques for Cancer Prediction: A Survey. *Procedia Computer Science, 60*, 769-774.

Ariyasu, S., Sawa, A., Morita, A., Hanaya, K., Hoshi, M., & Aoki, S. (2014). Design and Synthesis of 8-Hydroxyquinolin Based Radioprotective Agents. *Bioorganic and Medical Chemistry*, 22, 3891 - 3905.

Bekkar, M., Alitouche, T. A., & Djemaa, H. K. (2013). *Evaluation Measures for Models Assesment Over Imbalanced Data Sets*. *Journal of Information Engineering and Applications*, 3(10), 2224-5782.

Bishop, C. M. (2006). *Pattern Recognition And Machine Learning.* Cambridge: Springer.

Breiman, L. (2001). Random Forest. *Machine Learning*, 45(1), 5-32.

Breiman, L. F. (1993). *Classification and Regression Tree.* New York: Chapman Hall.

Chen, C., Liaw, A., & Breiman, L. (2004). *Using random forests to learn imbalanced.* California: *Statistics Department of University of California.*

Devi, M. A., Ravi, S., Vaishnavi, J., & Punitha, S. (2016). Classification of Cervical Cancer using Artificial Neural Networks. *Twelfth International Multi-Conference on Information Processing*, 89, 465 – 472.

Han, J., Kember, M., & Pei, J. (2012). *Data Mining Concepts and Technique 3rd Edition.* Waltham: Elsevier Inc.

Hong, H. W., & Sung, B. C. (2003). Paired neural network with negatively correlated features for cancer classification in DNA. *In Proceedings of the International Joint Conference on Neural Networks*, 1708 – 1713.

Hosmer, D. W. (2000). *Applied Logistic Regression (2nd ed.).* New York: John Wiley & Sons.

Jong, K. M., & Ki, K. D. (2010). Ensemble with neural networks for bankruptcy prediction. *Expert Systems with Applications*, 37, 3373–3379.

Kuswanto, H., Asfihani, A., Sarumaha, Y., & Ohwada, H. (2015). Logistic Regression Ensembles for Predicting Customer Defection with Very Large Sample Size. *Procedia Computer Science, 72*, 86-93.

Matohashi, H., Teraoka, T., Ohwada, H., & Aoki, S. (2018). *Regression Models and Ranking Method for P53 Inhibiotr Candidates Using Machine Learning*. *Bioorganic and Medicinal Chemistry*, 3891-3905.

NationalCancerInstitute. (2017), *National Cancer Institute*. diakses pada tangga 6 Februari 2019. *https://www.cancer.gov/about-cancer/treatment/types*. Retrieved from https://www.cancer.gov/.

Pappu, V., & Pardalos, P. M. (2014). *High Dimensional Data Classification*. *Clusters, Orders, and Trees: Methods and Applications: In Honor of Boris Mirkin's 70th Birthday*, 120-154.

Pratiwi, E. (2016). Analisis Dosis Radiasi Pada Radioterapi Eksternal *Linear Accelerator (Linac) Carcinoma Cervix* Dengan Variasi Sudut *Wedge* Di Rumah Sakit Umum Pusat Dr. Sardjito Yogyakarta. Universitas Gadjah Mada, Departemen Teknik Nuklir dan Teknik Fisika. Yogyakarta.

RI, K. K. (2013). *Riset Kesehatan Dasar*. Jakarta: Badan Litbang Kementerian Kesehatan RI.

Rusyidi, I. (2009). *Deteksi Dini dan Pencegahan Kanker Pada Wanita*. Jakarta: Sagung Seto.

Sarle, W. S. (1994). Neural Networks and Statistical Models. *Proceedings of the Nineteenth Annual SAS Users Group International Conference* (pp. 1538-1550). USA: SAS Institute.

Suhartono. (2007). *Feedforward Neural Network* untuk Pemodelan Runtun Waktu. Yogyakarta: Disertasi Doktoral, Universitas Gajah Mada.

TimCancerHelps. (2010). *Stop Kanker: Panduan Deteksi Dini & Pengobatan Menyeluruh Berbagai Jenis Kanker*. Jakarta Selatan: Agro Media Pustaka.

Wahyuni, T. (2015). Hubungan Antara Frekuensi Kemoterapi Dengan Kualitas Hidup Perempuan Dengan Kanker Payudara Yang Menjalani Kemoterapi Di Ruang Kemoterapi Rsud. A.M Parikesit Tenggarong. *Jurnal Ilmu Kesehatan, 3*, 1-3.

Walpole, R. E. (1995). *Pengantar Metode Statistika*. Diterjemahkan oleh Ir. Bambang Sumantri. Edisi Ketiga. Jakarta: Gramedia Pustaka Utama.

WHO. (2008). *Cancer*. Retrieved from Fact Sheets: http://www.who.int/mediacentre/factsheets/fs297/en/index.html

Wirasuta, I. A. (2006). *Buku Ajar : Toksologi Umum*. Bali: Universitas Udayana.

Zakharov, R., & Dupont, P. (2011). Ensemble logistic. *International Conference on Pattern Recognition* (pp. 133-144). Berlin: Springer Science.

*(Halaman ini sengaja dikosongkan)*

# LAMPIRAN

**Lampiran 1.** Data Citotoxity

| Obs | Name | X1 | X2 | ... | X216 | X217 | Kelas Aktual |
|---|---|---|---|---|---|---|---|
| | | pKa (max20) | Br Count | ... | Molecular 3D SAVol | Molecular Volume | |
| 1 | AS-11 | 7.9 | 0 | ... | 368.299 | 170.81 | 1 |
| 2 | AS-15 | 20 | 0 | ... | 417.73 | 202.36 | 0 |
| 3 | AS-16 | 10.1 | 0 | ... | 505.647 | 261.02 | 0 |
| 4 | AS-17 | 20 | 0 | ... | 530.977 | 288.11 | 0 |
| 5 | AS-2 | 11.5 | 0 | ... | 487.739 | 244.9 | 0 |
| 6 | KH-20 | 20 | 1 | ... | 468.618 | 239.41 | 0 |
| 7 | KH-24 | 10.5 | 0 | ... | 715.676 | 377.64 | 0 |
| 8 | KT-2 | 20 | 0 | ... | 648.995 | 333.73 | 0 |
| 9 | MH-16 | 11.8 | 0 | ... | 322.545 | 135.48 | 0 |
| 10 | MH-8 | 20 | 0 | ... | 505.138 | 274.05 | 0 |
| 11 | SAr-2 | 10.1 | 0 | ... | 445.044 | 209.91 | 0 |
| 12 | SAr-3 | 20 | 0 | ... | 562.259 | 286.06 | 0 |
| 13 | SAr-4 | 4.4 | 0 | ... | 396.157 | 173.21 | 0 |
| 14 | SAr-7 | 20 | 0 | ... | 912.486 | 487.74 | 1 |
| 15 | YM-13 | 20 | 1 | ... | 509.638 | 266.16 | 0 |
| 16 | YN-1 | 20 | 0 | ... | 808.732 | 428.4 | 0 |
| 17 | YN-7 | 20 | 0 | ... | 565.538 | 315.21 | 0 |
| 18 | YN-9 | 20 | 0 | ... | 566.256 | 306.98 | 0 |
| 19 | YT-1 | 9.8 | 0 | ... | 622.967 | 363.23 | 0 |
| 20 | AS-6 | 20 | 0 | ... | 460.551 | 234.95 | 0 |

**Lampiran 2.** Data Radioprotective

| Obs | Name | X1 | X2 | ... | X216 | X217 | Kelas Aktual |
|---|---|---|---|---|---|---|---|
| | | pKa (max20) | Br Count | ... | Molecular 3D SAVol | Molecular Volume | |
| 1 | AS-11 | 7.9 | 0 | ... | 368.299 | 170.81 | 0 |
| 2 | AS-15 | 20 | 0 | ... | 417.73 | 202.36 | 1 |
| 3 | AS-16 | 10.1 | 0 | ... | 505.647 | 261.02 | 1 |
| 4 | AS-17 | 20 | 0 | ... | 530.977 | 288.11 | 1 |
| 5 | AS-2 | 11.5 | 0 | ... | 487.739 | 244.9 | 1 |
| 6 | KH-20 | 20 | 1 | ... | 468.618 | 239.41 | 1 |
| 7 | KH-24 | 10.5 | 0 | ... | 715.676 | 377.64 | 1 |
| 8 | KT-2 | 20 | 0 | ... | 648.995 | 333.73 | 1 |
| 9 | MH-16 | 11.8 | 0 | ... | 322.545 | 135.48 | 1 |
| 10 | MH-8 | 20 | 0 | ... | 505.138 | 274.05 | 1 |
| 11 | SAr-2 | 10.1 | 0 | ... | 445.044 | 209.91 | 0 |
| 12 | SAr-3 | 20 | 0 | ... | 562.259 | 286.06 | 1 |
| 13 | SAr-4 | 4.4 | 0 | ... | 396.157 | 173.21 | 0 |
| 14 | SAr-7 | 20 | 0 | ... | 912.486 | 487.74 | 0 |
| 15 | YM-13 | 20 | 1 | ... | 509.638 | 266.16 | 1 |
| 16 | YN-1 | 20 | 0 | ... | 808.732 | 428.4 | 1 |
| 17 | YN-7 | 20 | 0 | ... | 565.538 | 315.21 | 0 |
| 18 | YN-9 | 20 | 0 | ... | 566.256 | 306.98 | 1 |
| 19 | YT-1 | 9.8 | 0 | ... | 622.967 | 363.23 | 1 |
| 20 | AS-6 | 20 | 0 | ... | 460.551 | 234.95 | 1 |

**Lampiran 3.** Syntax Random Forest

```
install.packages("randomForest")
library(randomForest)
#Input Data
mydata <- read.csv('D:/DATA.csv'. header = T)
data.X <- mydata[.-ncol(mydata)]
data.Y <- mydata[.ncol(mydata)]
#Random Forest Function
rforest <- function(y.x.fixthres = NULL) {
num_pred <- ncol(x)
num_obs <- nrow(x)
pos_rate <- sum(y)/num_obs
# thrashold search or default option
if(is.null(fixthres)) fixthres <- (pos_rate + .5)/2
#Estimation Random Forest
rf.model <- randomForest(x. y. importance = TRUE)
pred <- apply(data.frame(rf.model$predicted). 1. function(x)
{ifelse(x > fixthres. 1. 0)})
return(list(model = rf.model.treshold = fixthres.probability =
rf.model$predicted.predicted = pred))}
#Proses Running Random Forest
model.rf <- rforest(x = data.X.y = data.Y)
model.rf$treshold
model.rf$probability
confusionMatrix(as.factor(model.rf$predicted).as.factor(data.Y).p
ositive = "1")
plot(model.rf$model. type="l")
```

**Lampiran 4.** Syntax ANN *Ensemble*

```
install.packages("caret")
install.packages("https://github.com/jeremiedb/mxnet_winbin/raw
/master/mxnet.zip". repos = NULL)
library(caret)
library(mxnet)
#Input Data
mydata <- read.csv('D:/cito.csv'. header = T)
data.X <- mydata[.-ncol(mydata)]
data.Y <- mydata[.ncol(mydata)]

#Neural Net Ensemble Function
NN.cerp <-
function(y.x.nens.facv.fixsize=NULL.fixthres=NULL) {
# initialization
set.seed(as.numeric(Sys.time()))
options(warn=-1)
if(sum(is.na(x))>0) stop("missing value is found")
if(sum(is.na(y))>0) stop("missing value is found")
num_pred <- ncol(x)
num_obs <- nrow(x)
pos_rate <- sum(y)/num_obs
# parameter search or default option
if(is.null(fixthres)) fixthres <- (pos_rate + .5)/2
if(is.null(fixsize)) fixsize <- round(6*num_pred/num_obs)
# main body
if(facv == "softmax") {
   ptss <- floor(seq(1.fixsize+.999.length.out=num_pred))
   fitted <- NULL; cname <- NULL; partition.table <-
matrix(0.num_pred.nens)
   for (i in 1:nens) {
     cname <- c(cname. paste("ens".i.sep=""))
rand_pred <- sample(ptss)
     partition.table[.i] <- rand_pred
```

**Lampiran 4.** Syntax ANN *Ensemble* (lanjutan)

```
avg_fit <- rep(0.num_obs)
  for(j in 1:fixsize) {
        smp_dt <- cbind(x[.rand_pred==j])
        ANNmodel <- mx.mlp(data = as.matrix(smp_dt)
rand_pred <- sample(ptss)
      partition.table[.i] <- rand_pred
      avg_fit <- rep(0.num_obs)
      for(j in 1:fixsize) {
rand_pred <- sample(ptss)
      partition.table[.i] <- rand_pred
      avg_fit <- rep(0.num_obs)
      for(j in 1:fixsize) {
        smp_dt <- cbind(x[.rand_pred==j])
        ANNmodel <- mx.mlp(data = as.matrix(smp_dt)
                           .label = y
                           .hidden_node = c(ncol(smp_dt)) #A layers - first
layer
                           .out_node = 2 #number of output nodes
                           .out_activation = facv #activation function
                           .num.round = 10 #number of iterations
                           .array.batch.size = 5 #batch size for updating
weights
                           .learning.rate = 0.05 #same as step size
                           .eval.metric= mx.metric.accuracy #matrix
evaluation
                           .array.layout = "rowmajor")

        preds <- predict(ANNmodel. as.matrix(smp_dt).array.layout
= "rowmajor")[2.]
        avg_fit <- avg_fit + preds
      }
      fitted <- cbind(fitted.avg_fit/fixsize.deparse.level = 0)
    }
```

**Lampiran 4.** Syntax ANN *Ensemble* (lanjutan)

```
  }
  else {
    ptss <- floor(seq(1.fixsize+.999.length.out=num_pred))
    fitted <- NULL; cname <- NULL; partition.table <-
matrix(0.num_pred.nens)
    for (i in 1:nens) {
cname <- c(cname. paste("ens".i.sep=""))
      rand_pred <- sample(ptss)
      partition.table[.i] <- rand_pred
      avg_fit <- rep(0.num_obs)
      for(j in 1:fixsize) {
     smp_dt <- cbind(x[.rand_pred==j])
        ANNmodel <- mx.mlp(data = as.matrix(smp_dt)
                          .label = y
                          .hidden_node = c(ncol(smp_dt)) #A layers - first
layer
                          .out_node = 1 #number of output nodes
                          .out_activation = facv #activation function
                          .num.round = 10 #number of iterations
                          .array.batch.size = 5 #batch size for updating
weights
                          .learning.rate = 0.05 #same as step size
                          .eval.metric= mx.metric.accuracy #matrix
evaluation
                          .array.layout = "rowmajor")

        preds <- predict(ANNmodel. as.matrix(smp_dt).array.layout
= "rowmajor")[1.]
        avg_fit <- avg_fit + preds
      }
      fitted <- cbind(fitted.avg_fit/fixsize.deparse.level = 0)
    }
probability <- rowMeans(fitted)
```

**Lampiran 4.** Syntax ANN *Ensemble* (lanjutan)

```
learning.decision <- apply(cbind(probability). 1. function(x)
{ifelse(x > fixthres. 1. 0)})
  colnames(fitted) <- cname
colnames(partition.table) <- cname; rownames(partition.table) <-
colnames(x)
return(list(probability.ens=fitted.probability.final=probability.lear
ning.decision=learning.decision.
partition.table=partition.table.number.ensemble=nens.optimal.siz
e=fixsize.optimal.threshold=fixthres))
}
#Proses Running ANN Ensemble with softmax activation
function
colnames(partition.table) <- cname; rownames(partition.table) <-
colnames(x)
return(list(probability.ens=fitted.probability.final=probability.lear
ning.decision=learning.decision.

partition.table=partition.table.number.ensemble=nens.optimal.siz
e=fixsize.optimal.threshold=fixthres))
}
#Proses Running ANN Ensemble with softmax activation
function
model.annA <- NN.cerp(y = data.Y.x = data.X.nens = 11.facv =
"softmax".fixsize = A)
model.annA$partition.table
model.annA$optimal.threshold
model.annA$probability.ens
model.annA$probability.final
model.annA$learning.decision
confusionMatrix(as.factor(model.annA$learning.decision).as.fact
or(data.Y).positive = "1")
#Proses Running ANN Ensemble with sigmoid activation
function
```

**Lampiran 4.** Syntax ANN *Ensemble* (lanjutan)

```
model.annA<- NN.cerp(y = data.Y.x = data.X.nens = 11.facv =
"logistic".fixsize = A)
model.annA$partition.table
model.annA$optimal.threshold
model.annA$probability.ens
model.annA$probability.final
model.annA$learning.decision
confusionMatrix(as.factor(model.annA$learning.decision).as.fact
or(data.Y).positive = "1")
```

**Lampiran 5.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 12 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| x1 | 9 | 9 | 3 | 3 | 11 | 11 | 3 | 5 | 10 | 1 | 6 |
| x2 | 1 | 7 | 6 | 1 | 1 | 9 | 7 | 10 | 9 | 6 | 6 |
| x3 | 1 | 5 | 3 | 11 | 2 | 8 | 11 | 7 | 4 | 12 | 9 |
| x4 | 6 | 3 | 10 | 4 | 11 | 5 | 9 | 3 | 5 | 7 | 2 |
| x5 | 8 | 5 | 12 | 10 | 7 | 10 | 3 | 2 | 10 | 11 | 6 |
| x6 | 9 | 4 | 8 | 2 | 9 | 2 | 12 | 8 | 9 | 6 | 10 |
| x7 | 2 | 4 | 11 | 7 | 5 | 2 | 9 | 10 | 9 | 9 | 12 |
| x8 | 12 | 12 | 3 | 7 | 4 | 12 | 12 | 6 | 12 | 1 | 9 |
| x9 | 9 | 1 | 10 | 5 | 6 | 12 | 7 | 9 | 3 | 9 | 12 |
| x10 | 4 | 8 | 11 | 7 | 11 | 9 | 5 | 11 | 1 | 3 | 6 |
| x11 | 10 | 10 | 1 | 10 | 9 | 4 | 1 | 7 | 1 | 1 | 11 |
| x12 | 9 | 3 | 6 | 7 | 3 | 1 | 10 | 6 | 8 | 9 | 1 |
| x13 | 8 | 7 | 5 | 1 | 4 | 4 | 2 | 3 | 3 | 11 | 2 |
| x14 | 4 | 11 | 2 | 3 | 5 | 8 | 5 | 2 | 6 | 4 | 7 |
| x15 | 2 | 4 | 9 | 1 | 9 | 9 | 12 | 7 | 8 | 10 | 8 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| x84 | 11 | 2 | 3 | 3 | 1 | 8 | 7 | 7 | 5 | 6 | 3 |
| x85 | 4 | 5 | 5 | 10 | 2 | 6 | 9 | 6 | 4 | 11 | 8 |
| x86 | 10 | 11 | 4 | 5 | 1 | 10 | 9 | 9 | 6 | 5 | 11 |
| x87 | 4 | 2 | 6 | 11 | 3 | 4 | 1 | 5 | 11 | 6 | 1 |
| x88 | 6 | 11 | 11 | 3 | 12 | 9 | 3 | 10 | 2 | 5 | 5 |
| x89 | 10 | 1 | 12 | 11 | 10 | 4 | 2 | 1 | 5 | 10 | 1 |
| x90 | 7 | 11 | 11 | 12 | 12 | 6 | 8 | 12 | 9 | 1 | 1 |

**Lampiran 6.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 13 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| x1 | 10 | 3 | 6 | 11 | 3 | 1 | 2 | 7 | 6 | 12 | 10 |
| x2 | 8 | 2 | 5 | 7 | 4 | 6 | 6 | 5 | 5 | 8 | 8 |
| x3 | 5 | 9 | 3 | 5 | 6 | 7 | 5 | 8 | 12 | 6 | 4 |
| x4 | 6 | 13 | 7 | 7 | 9 | 2 | 5 | 5 | 9 | 11 | 10 |
| x5 | 12 | 2 | 7 | 11 | 5 | 11 | 5 | 13 | 9 | 13 | 9 |
| x6 | 1 | 7 | 1 | 7 | 9 | 13 | 2 | 11 | 7 | 1 | 9 |
| x7 | 5 | 10 | 10 | 10 | 11 | 1 | 11 | 7 | 4 | 13 | 10 |
| x8 | 6 | 8 | 9 | 12 | 11 | 5 | 10 | 7 | 10 | 13 | 12 |
| x9 | 4 | 3 | 5 | 9 | 7 | 13 | 9 | 12 | 6 | 9 | 1 |
| x10 | 4 | 1 | 4 | 8 | 9 | 2 | 9 | 4 | 7 | 5 | 9 |
| x11 | 2 | 5 | 11 | 7 | 13 | 6 | 3 | 4 | 5 | 5 | 8 |
| x12 | 1 | 4 | 3 | 4 | 2 | 10 | 9 | 9 | 6 | 13 | 2 |
| x13 | 3 | 7 | 11 | 7 | 4 | 12 | 4 | 12 | 6 | 11 | 2 |
| x14 | 3 | 8 | 13 | 6 | 12 | 7 | 10 | 2 | 10 | 11 | 1 |
| x15 | 13 | 5 | 2 | 8 | 12 | 4 | 6 | 1 | 7 | 13 | 10 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| x84 | 8 | 7 | 10 | 3 | 13 | 10 | 12 | 1 | 2 | 8 | 13 |
| x85 | 7 | 3 | 1 | 3 | 2 | 3 | 4 | 7 | 12 | 12 | 8 |
| x86 | 12 | 9 | 10 | 9 | 12 | 3 | 2 | 6 | 6 | 3 | 1 |
| x87 | 9 | 1 | 7 | 1 | 9 | 8 | 4 | 11 | 6 | 5 | 4 |
| x88 | 7 | 12 | 7 | 10 | 4 | 2 | 2 | 5 | 9 | 1 | 9 |
| x89 | 3 | 5 | 6 | 4 | 9 | 10 | 2 | 3 | 8 | 9 | 13 |
| x90 | 4 | 10 | 1 | 10 | 3 | 7 | 9 | 2 | 13 | 11 | 13 |

**Lampiran 7.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 14 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 14 | 10 | 14 | 13 | 13 | 12 | 7 | 11 | 14 | 6 | 3 |
| X2 | 6 | 9 | 9 | 13 | 1 | 3 | 13 | 2 | 14 | 5 | 4 |
| X3 | 12 | 2 | 2 | 1 | 3 | 14 | 2 | 10 | 10 | 9 | 5 |
| X4 | 4 | 9 | 12 | 3 | 14 | 3 | 6 | 14 | 14 | 7 | 14 |
| X5 | 14 | 3 | 5 | 7 | 8 | 9 | 14 | 3 | 11 | 13 | 13 |
| X6 | 9 | 8 | 4 | 12 | 8 | 4 | 11 | 1 | 9 | 5 | 8 |
| X7 | 10 | 1 | 9 | 1 | 9 | 10 | 5 | 2 | 4 | 10 | 14 |
| X8 | 3 | 13 | 8 | 8 | 4 | 10 | 5 | 7 | 9 | 1 | 11 |
| X9 | 2 | 7 | 8 | 14 | 10 | 9 | 12 | 11 | 1 | 7 | 4 |
| X10 | 6 | 12 | 9 | 2 | 4 | 1 | 9 | 5 | 2 | 12 | 10 |
| X11 | 2 | 2 | 12 | 11 | 9 | 9 | 10 | 6 | 10 | 4 | 9 |
| X12 | 3 | 8 | 3 | 8 | 6 | 13 | 8 | 4 | 1 | 7 | 10 |
| X13 | 3 | 13 | 5 | 7 | 1 | 2 | 7 | 6 | 2 | 6 | 7 |
| X14 | 3 | 8 | 9 | 7 | 5 | 1 | 13 | 5 | 1 | 11 | 12 |
| X15 | 4 | 3 | 10 | 2 | 2 | 3 | 7 | 10 | 8 | 5 | 5 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 8 | 5 | 6 | 10 | 12 | 2 | 2 | 9 | 1 | 7 | 5 |
| X85 | 5 | 4 | 6 | 9 | 5 | 14 | 14 | 13 | 4 | 14 | 5 |
| X86 | 12 | 3 | 3 | 1 | 4 | 1 | 14 | 11 | 12 | 4 | 14 |
| X87 | 3 | 9 | 8 | 9 | 4 | 6 | 13 | 7 | 12 | 12 | 4 |
| X88 | 6 | 10 | 14 | 1 | 9 | 2 | 13 | 8 | 6 | 14 | 13 |
| X89 | 8 | 3 | 13 | 8 | 12 | 6 | 1 | 13 | 5 | 14 | 10 |
| X90 | 9 | 5 | 7 | 5 | 7 | 8 | 12 | 8 | 2 | 7 | 14 |

**Lampiran 8.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 15 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax*

| **Compound** | ***Ensemble*** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| X1 | 4 | 8 | 11 | 11 | 2 | 6 | 11 | 2 | 7 | 4 | 14 |
| X2 | 7 | 6 | 9 | 12 | 10 | 3 | 9 | 4 | 7 | 14 | 10 |
| X3 | 9 | 2 | 11 | 7 | 4 | 14 | 7 | 7 | 1 | 1 | 11 |
| X4 | 9 | 12 | 7 | 5 | 6 | 6 | 13 | 14 | 10 | 5 | 1 |
| X5 | 11 | 10 | 8 | 12 | 2 | 14 | 9 | 14 | 11 | 12 | 8 |
| X6 | 7 | 10 | 14 | 5 | 4 | 6 | 14 | 7 | 8 | 12 | 12 |
| X7 | 2 | 12 | 11 | 5 | 15 | 9 | 5 | 13 | 14 | 3 | 7 |
| X8 | 5 | 2 | 2 | 3 | 10 | 8 | 4 | 6 | 13 | 10 | 1 |
| X9 | 3 | 13 | 5 | 13 | 7 | 10 | 1 | 10 | 9 | 1 | 13 |
| X10 | 11 | 15 | 5 | 2 | 3 | 8 | 10 | 5 | 14 | 9 | 10 |
| X11 | 9 | 8 | 4 | 9 | 4 | 9 | 15 | 13 | 6 | 5 | 1 |
| X12 | 15 | 15 | 11 | 1 | 1 | 4 | 10 | 12 | 12 | 4 | 6 |
| X13 | 7 | 6 | 13 | 13 | 10 | 14 | 8 | 15 | 8 | 10 | 11 |
| X14 | 12 | 10 | 10 | 14 | 4 | 3 | 12 | 6 | 11 | 1 | 5 |
| X15 | 13 | 4 | 15 | 4 | 8 | 11 | 2 | 4 | 7 | 2 | 13 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 5 | 12 | 12 | 6 | 6 | 12 | 10 | 9 | 9 | 1 | 14 |
| X85 | 3 | 10 | 1 | 15 | 9 | 7 | 1 | 9 | 10 | 11 | 4 |
| X86 | 2 | 11 | 4 | 14 | 12 | 6 | 10 | 10 | 10 | 12 | 7 |
| X87 | 1 | 9 | 14 | 12 | 6 | 11 | 11 | 15 | 7 | 7 | 3 |
| X88 | 12 | 11 | 4 | 15 | 14 | 1 | 6 | 6 | 7 | 8 | 12 |
| X89 | 3 | 8 | 5 | 12 | 3 | 5 | 15 | 13 | 5 | 13 | 4 |
| X90 | 8 | 15 | 4 | 3 | 4 | 2 | 9 | 3 | 10 | 8 | 12 |

**Lampiran 9.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 16 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax*

| **Compound** | ***Ensemble*** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| X1 | 7 | 6 | 1 | 12 | 8 | 2 | 16 | 8 | 6 | 14 | 13 |
| X2 | 13 | 16 | 11 | 16 | 11 | 1 | 1 | 4 | 13 | 10 | 15 |
| X3 | 7 | 10 | 1 | 5 | 4 | 9 | 16 | 3 | 11 | 5 | 13 |
| X4 | 5 | 11 | 7 | 14 | 14 | 10 | 1 | 10 | 9 | 9 | 1 |
| X5 | 1 | 16 | 5 | 16 | 16 | 2 | 14 | 10 | 10 | 16 | 11 |
| X6 | 3 | 4 | 12 | 10 | 2 | 9 | 9 | 5 | 8 | 15 | 1 |
| X7 | 13 | 13 | 7 | 8 | 3 | 15 | 5 | 2 | 10 | 7 | 14 |
| X8 | 13 | 13 | 8 | 12 | 15 | 16 | 7 | 1 | 7 | 4 | 11 |
| X9 | 16 | 2 | 3 | 4 | 14 | 14 | 11 | 14 | 2 | 5 | 10 |
| X10 | 13 | 2 | 6 | 15 | 6 | 3 | 3 | 16 | 7 | 2 | 8 |
| X11 | 9 | 1 | 13 | 7 | 7 | 10 | 12 | 12 | 3 | 16 | 13 |
| X12 | 16 | 2 | 5 | 7 | 14 | 4 | 2 | 4 | 4 | 12 | 9 |
| X13 | 12 | 1 | 4 | 8 | 11 | 9 | 6 | 8 | 2 | 1 | 3 |
| X14 | 1 | 4 | 12 | 9 | 16 | 6 | 11 | 11 | 15 | 6 | 16 |
| X15 | 10 | 11 | 2 | 16 | 10 | 1 | 15 | 5 | 5 | 8 | 6 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 12 | 7 | 8 | 12 | 2 | 10 | 9 | 8 | 6 | 14 | 5 |
| X85 | 6 | 15 | 16 | 11 | 6 | 9 | 7 | 14 | 16 | 3 | 4 |
| X86 | 5 | 5 | 15 | 4 | 12 | 7 | 6 | 5 | 6 | 8 | 10 |
| X87 | 4 | 12 | 15 | 5 | 15 | 14 | 3 | 15 | 1 | 10 | 12 |
| X88 | 12 | 13 | 13 | 2 | 12 | 9 | 13 | 10 | 3 | 4 | 12 |
| X89 | 14 | 11 | 16 | 3 | 9 | 11 | 14 | 2 | 3 | 2 | 6 |
| X90 | 4 | 9 | 16 | 12 | 10 | 3 | 7 | 6 | 8 | 6 | 6 |

**Lampiran 10.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 17 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| X1 | 13 | 4 | 13 | 8 | 14 | 16 | 7 | 17 | 2 | 10 | 10 |
| X2 | 12 | 4 | 15 | 4 | 9 | 14 | 17 | 9 | 7 | 16 | 17 |
| X3 | 17 | 3 | 5 | 7 | 10 | 8 | 4 | 16 | 9 | 2 | 11 |
| X4 | 13 | 6 | 13 | 11 | 2 | 15 | 10 | 2 | 15 | 9 | 17 |
| X5 | 6 | 5 | 15 | 4 | 13 | 15 | 16 | 5 | 2 | 1 | 3 |
| X6 | 7 | 15 | 9 | 10 | 17 | 1 | 5 | 1 | 1 | 1 | 12 |
| X7 | 7 | 3 | 8 | 10 | 15 | 3 | 1 | 4 | 7 | 17 | 6 |
| X8 | 10 | 9 | 8 | 11 | 8 | 1 | 6 | 12 | 7 | 4 | 8 |
| X9 | 15 | 1 | 1 | 13 | 5 | 2 | 11 | 17 | 12 | 4 | 2 |
| X10 | 16 | 17 | 6 | 9 | 5 | 9 | 4 | 11 | 11 | 15 | 9 |
| X11 | 5 | 8 | 10 | 14 | 15 | 17 | 3 | 11 | 16 | 14 | 17 |
| X12 | 15 | 6 | 3 | 8 | 3 | 17 | 1 | 13 | 14 | 10 | 4 |
| X13 | 6 | 9 | 13 | 1 | 11 | 16 | 5 | 5 | 9 | 6 | 10 |
| X14 | 16 | 9 | 3 | 6 | 14 | 6 | 6 | 5 | 4 | 7 | 1 |
| X15 | 4 | 15 | 10 | 12 | 11 | 17 | 17 | 14 | 6 | 12 | 17 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 3 | 1 | 12 | 12 | 7 | 11 | 8 | 15 | 5 | 11 | 1 |
| X85 | 11 | 5 | 17 | 1 | 9 | 8 | 10 | 17 | 17 | 2 | 5 |
| X86 | 2 | 16 | 5 | 6 | 9 | 2 | 13 | 17 | 1 | 1 | 16 |
| X87 | 15 | 9 | 8 | 14 | 9 | 13 | 7 | 9 | 15 | 13 | 3 |
| X88 | 11 | 17 | 11 | 14 | 15 | 16 | 1 | 4 | 4 | 13 | 13 |
| X89 | 10 | 7 | 3 | 12 | 6 | 4 | 3 | 12 | 7 | 15 | 11 |
| X90 | 7 | 5 | 12 | 9 | 12 | 1 | 13 | 13 | 4 | 10 | 15 |

**Lampiran 11.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 18 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 15 | 18 | 3 | 16 | 8 | 14 | 17 | 7 | 15 | 18 | 10 |
| X2 | 15 | 11 | 8 | 1 | 5 | 1 | 11 | 14 | 5 | 10 | 18 |
| X3 | 3 | 3 | 15 | 5 | 5 | 6 | 17 | 13 | 15 | 15 | 18 |
| X4 | 7 | 15 | 16 | 1 | 6 | 9 | 10 | 7 | 11 | 6 | 4 |
| X5 | 2 | 17 | 2 | 13 | 18 | 7 | 9 | 3 | 10 | 9 | 18 |
| X6 | 11 | 6 | 16 | 12 | 6 | 10 | 6 | 9 | 7 | 7 | 11 |
| X7 | 13 | 9 | 13 | 3 | 18 | 15 | 9 | 14 | 10 | 14 | 9 |
| X8 | 10 | 1 | 9 | 10 | 3 | 15 | 6 | 14 | 12 | 14 | 6 |
| X9 | 9 | 9 | 11 | 6 | 9 | 8 | 2 | 16 | 8 | 2 | 12 |
| X10 | 18 | 13 | 12 | 6 | 2 | 7 | 8 | 10 | 18 | 15 | 9 |
| X11 | 11 | 6 | 6 | 14 | 10 | 10 | 13 | 9 | 1 | 14 | 16 |
| X12 | 2 | 18 | 13 | 13 | 5 | 16 | 18 | 7 | 2 | 1 | 6 |
| X13 | 10 | 15 | 4 | 18 | 3 | 17 | 4 | 3 | 18 | 16 | 15 |
| X14 | 14 | 11 | 9 | 17 | 4 | 10 | 9 | 1 | 3 | 17 | 3 |
| X15 | 4 | 16 | 6 | 18 | 12 | 1 | 14 | 12 | 4 | 10 | 18 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 15 | 10 | 5 | 13 | 15 | 13 | 15 | 16 | 3 | 16 | 1 |
| X85 | 11 | 16 | 11 | 5 | 18 | 15 | 18 | 17 | 2 | 13 | 15 |
| X86 | 17 | 4 | 10 | 1 | 12 | 6 | 16 | 14 | 9 | 16 | 15 |
| X87 | 17 | 14 | 5 | 8 | 17 | 18 | 9 | 2 | 15 | 5 | 7 |
| X88 | 11 | 16 | 13 | 5 | 1 | 18 | 10 | 7 | 14 | 15 | 7 |
| X89 | 11 | 14 | 16 | 9 | 13 | 10 | 12 | 9 | 1 | 12 | 7 |
| X90 | 10 | 17 | 14 | 14 | 3 | 9 | 3 | 15 | 2 | 3 | 1 |

**Lampiran 12.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 19 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Softmax*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 9 | 3 | 1 | 8 | 6 | 6 | 14 | 17 | 19 | 11 | 13 |
| X2 | 16 | 5 | 19 | 12 | 12 | 16 | 16 | 11 | 12 | 13 | 14 |
| X3 | 8 | 6 | 6 | 4 | 11 | 10 | 14 | 12 | 10 | 13 | 3 |
| X4 | 15 | 15 | 13 | 19 | 5 | 16 | 11 | 16 | 14 | 6 | 5 |
| X5 | 19 | 6 | 11 | 10 | 18 | 19 | 13 | 14 | 9 | 18 | 16 |
| X6 | 14 | 7 | 15 | 11 | 15 | 14 | 19 | 4 | 3 | 10 | 12 |
| X7 | 5 | 14 | 17 | 8 | 8 | 7 | 17 | 11 | 17 | 7 | 4 |
| X8 | 9 | 1 | 10 | 17 | 4 | 11 | 5 | 3 | 18 | 11 | 4 |
| X9 | 18 | 11 | 6 | 6 | 15 | 13 | 5 | 2 | 16 | 8 | 17 |
| X10 | 18 | 17 | 19 | 6 | 9 | 7 | 7 | 9 | 11 | 11 | 2 |
| X11 | 7 | 18 | 13 | 2 | 1 | 16 | 6 | 15 | 6 | 16 | 3 |
| X12 | 18 | 15 | 18 | 14 | 16 | 17 | 2 | 4 | 11 | 5 | 4 |
| X13 | 6 | 4 | 16 | 13 | 5 | 4 | 16 | 5 | 7 | 7 | 9 |
| X14 | 19 | 16 | 15 | 11 | 7 | 16 | 1 | 6 | 12 | 14 | 8 |
| X15 | 13 | 9 | 5 | 7 | 12 | 7 | 6 | 10 | 2 | 4 | 3 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 10 | 10 | 14 | 4 | 16 | 12 | 12 | 1 | 9 | 16 | 16 |
| X85 | 3 | 9 | 9 | 11 | 3 | 9 | 19 | 6 | 10 | 8 | 3 |
| X86 | 6 | 2 | 17 | 2 | 7 | 8 | 7 | 14 | 12 | 3 | 10 |
| X87 | 18 | 16 | 15 | 9 | 6 | 8 | 18 | 18 | 8 | 17 | 4 |
| X88 | 4 | 14 | 18 | 19 | 19 | 2 | 2 | 1 | 17 | 4 | 14 |
| X89 | 6 | 8 | 3 | 9 | 5 | 16 | 1 | 13 | 17 | 18 | 7 |
| X90 | 16 | 3 | 18 | 5 | 9 | 9 | 4 | 1 | 11 | 3 | 14 |

**Lampiran 13.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 12 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*

| **Compound** | ***Ensemble*** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| X1 | 3 | 12 | 8 | 5 | 4 | 4 | 1 | 6 | 1 | 6 | 6 |
| X2 | 9 | 3 | 4 | 5 | 12 | 12 | 10 | 2 | 6 | 8 | 2 |
| X3 | 8 | 1 | 6 | 8 | 5 | 7 | 10 | 9 | 4 | 2 | 7 |
| X4 | 12 | 7 | 6 | 7 | 5 | 3 | 7 | 5 | 9 | 7 | 11 |
| X5 | 12 | 1 | 5 | 4 | 1 | 7 | 11 | 4 | 11 | 3 | 8 |
| X6 | 5 | 3 | 9 | 6 | 5 | 3 | 3 | 1 | 4 | 5 | 2 |
| X7 | 6 | 6 | 11 | 9 | 6 | 1 | 11 | 12 | 1 | 11 | 7 |
| X8 | 8 | 9 | 6 | 9 | 11 | 1 | 1 | 12 | 1 | 7 | 4 |
| X9 | 9 | 10 | 10 | 11 | 7 | 6 | 2 | 12 | 8 | 2 | 1 |
| X10 | 8 | 3 | 10 | 4 | 4 | 8 | 10 | 7 | 9 | 2 | 7 |
| X11 | 1 | 11 | 8 | 8 | 2 | 2 | 7 | 3 | 10 | 1 | 3 |
| X12 | 12 | 2 | 1 | 8 | 5 | 8 | 10 | 3 | 8 | 12 | 10 |
| X13 | 5 | 4 | 3 | 12 | 7 | 2 | 6 | 7 | 10 | 1 | 6 |
| X14 | 1 | 12 | 12 | 11 | 12 | 2 | 12 | 8 | 4 | 5 | 3 |
| X15 | 12 | 9 | 1 | 1 | 10 | 12 | 12 | 7 | 10 | 5 | 1 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 3 | 6 | 7 | 4 | 4 | 4 | 12 | 11 | 11 | 3 | 8 |
| X85 | 6 | 4 | 2 | 2 | 2 | 10 | 5 | 3 | 10 | 12 | 5 |
| X86 | 11 | 11 | 12 | 9 | 8 | 11 | 5 | 10 | 7 | 1 | 10 |
| X87 | 4 | 9 | 3 | 1 | 9 | 5 | 4 | 2 | 5 | 1 | 3 |
| X88 | 9 | 12 | 3 | 10 | 8 | 5 | 6 | 2 | 2 | 2 | 3 |
| X89 | 2 | 2 | 11 | 9 | 7 | 2 | 9 | 10 | 4 | 9 | 12 |
| X90 | 5 | 1 | 7 | 3 | 2 | 10 | 8 | 9 | 11 | 7 | 2 |

**Lampiran 14.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 13 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*

| **Compound** | ***Ensemble*** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| X1 | 10 | 3 | 6 | 11 | 3 | 1 | 2 | 7 | 6 | 12 | 10 |
| X2 | 8 | 2 | 5 | 7 | 4 | 6 | 6 | 5 | 5 | 8 | 8 |
| X3 | 5 | 9 | 3 | 5 | 6 | 7 | 5 | 8 | 12 | 6 | 4 |
| X4 | 6 | 13 | 7 | 7 | 9 | 2 | 5 | 5 | 9 | 11 | 10 |
| X5 | 12 | 2 | 7 | 11 | 5 | 11 | 5 | 13 | 9 | 13 | 9 |
| X6 | 1 | 7 | 1 | 7 | 9 | 13 | 2 | 11 | 7 | 1 | 9 |
| X7 | 5 | 10 | 10 | 10 | 11 | 1 | 11 | 7 | 4 | 13 | 10 |
| X8 | 6 | 8 | 9 | 12 | 11 | 5 | 10 | 7 | 10 | 13 | 12 |
| X9 | 4 | 3 | 5 | 9 | 7 | 13 | 9 | 12 | 6 | 9 | 1 |
| X10 | 4 | 1 | 4 | 8 | 9 | 2 | 9 | 4 | 7 | 5 | 9 |
| X11 | 2 | 5 | 11 | 7 | 13 | 6 | 3 | 4 | 5 | 5 | 8 |
| X12 | 1 | 4 | 3 | 4 | 2 | 10 | 9 | 9 | 6 | 13 | 2 |
| X13 | 3 | 7 | 11 | 7 | 4 | 12 | 4 | 12 | 6 | 11 | 2 |
| X14 | 3 | 8 | 13 | 6 | 12 | 7 | 10 | 2 | 10 | 11 | 1 |
| X15 | 13 | 5 | 2 | 8 | 12 | 4 | 6 | 1 | 7 | 13 | 10 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 8 | 7 | 10 | 3 | 13 | 10 | 12 | 1 | 2 | 8 | 13 |
| X85 | 7 | 3 | 1 | 3 | 2 | 3 | 4 | 7 | 12 | 12 | 8 |
| X86 | 12 | 9 | 10 | 9 | 12 | 3 | 2 | 6 | 6 | 3 | 1 |
| X87 | 9 | 1 | 7 | 1 | 9 | 8 | 4 | 11 | 6 | 5 | 4 |
| X88 | 7 | 12 | 7 | 10 | 4 | 2 | 2 | 5 | 9 | 1 | 9 |
| X89 | 3 | 5 | 6 | 4 | 9 | 10 | 2 | 3 | 8 | 9 | 13 |
| X90 | 4 | 10 | 1 | 10 | 3 | 7 | 9 | 2 | 13 | 11 | 13 |

**Lampiran 15.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 14 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 1 | 9 | 14 | 1 | 3 | 2 | 8 | 13 | 8 | 10 | 13 |
| X2 | 10 | 5 | 11 | 2 | 2 | 14 | 14 | 10 | 2 | 9 | 6 |
| X3 | 8 | 1 | 14 | 4 | 9 | 10 | 14 | 10 | 13 | 13 | 3 |
| X4 | 14 | 12 | 7 | 10 | 2 | 7 | 11 | 3 | 13 | 2 | 2 |
| X5 | 10 | 11 | 2 | 6 | 2 | 6 | 1 | 12 | 9 | 4 | 10 |
| X6 | 6 | 6 | 14 | 8 | 14 | 11 | 13 | 8 | 13 | 5 | 2 |
| X7 | 7 | 12 | 2 | 4 | 8 | 14 | 8 | 1 | 14 | 8 | 7 |
| X8 | 4 | 13 | 6 | 12 | 9 | 9 | 7 | 9 | 14 | 5 | 8 |
| X9 | 4 | 14 | 14 | 5 | 13 | 12 | 10 | 2 | 13 | 13 | 7 |
| X10 | 13 | 5 | 11 | 6 | 14 | 14 | 11 | 8 | 13 | 13 | 14 |
| X11 | 3 | 2 | 7 | 13 | 12 | 7 | 12 | 4 | 9 | 14 | 1 |
| X12 | 1 | 14 | 10 | 1 | 5 | 13 | 3 | 3 | 6 | 12 | 13 |
| X13 | 4 | 12 | 10 | 11 | 6 | 8 | 9 | 5 | 9 | 9 | 14 |
| X14 | 3 | 6 | 3 | 14 | 11 | 4 | 1 | 2 | 10 | 12 | 11 |
| X15 | 7 | 3 | 14 | 8 | 6 | 14 | 1 | 14 | 10 | 3 | 14 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 9 | 8 | 1 | 8 | 7 | 9 | 3 | 1 | 8 | 11 | 13 |
| X85 | 1 | 1 | 5 | 12 | 8 | 12 | 14 | 9 | 12 | 1 | 9 |
| X86 | 12 | 5 | 3 | 10 | 11 | 13 | 3 | 13 | 1 | 11 | 7 |
| X87 | 12 | 1 | 6 | 5 | 13 | 12 | 13 | 2 | 9 | 3 | 3 |
| X88 | 10 | 2 | 10 | 14 | 2 | 1 | 14 | 14 | 11 | 3 | 4 |
| X89 | 13 | 4 | 13 | 5 | 12 | 3 | 7 | 1 | 13 | 8 | 10 |
| X90 | 9 | 8 | 8 | 5 | 11 | 1 | 2 | 14 | 10 | 11 | 11 |

**Lampiran 16.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 15 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 2 | 10 | 3 | 13 | 10 | 15 | 15 | 11 | 7 | 5 | 2 |
| X2 | 5 | 1 | 2 | 11 | 4 | 7 | 12 | 8 | 3 | 5 | 13 |
| X3 | 6 | 13 | 6 | 13 | 9 | 5 | 3 | 9 | 6 | 12 | 8 |
| X4 | 14 | 12 | 4 | 5 | 11 | 5 | 1 | 3 | 4 | 1 | 1 |
| X5 | 1 | 5 | 7 | 9 | 10 | 14 | 15 | 3 | 13 | 4 | 1 |
| X6 | 12 | 9 | 14 | 10 | 8 | 11 | 12 | 10 | 9 | 8 | 15 |
| X7 | 8 | 11 | 6 | 13 | 15 | 15 | 15 | 11 | 3 | 9 | 6 |
| X8 | 8 | 12 | 4 | 14 | 15 | 7 | 1 | 14 | 11 | 5 | 7 |
| X9 | 3 | 9 | 5 | 2 | 12 | 10 | 4 | 10 | 10 | 14 | 10 |
| X10 | 5 | 5 | 14 | 8 | 10 | 5 | 2 | 15 | 7 | 12 | 6 |
| X11 | 6 | 8 | 2 | 5 | 14 | 8 | 15 | 1 | 9 | 8 | 14 |
| X12 | 10 | 1 | 9 | 2 | 2 | 11 | 7 | 4 | 12 | 7 | 1 |
| X13 | 15 | 1 | 15 | 12 | 3 | 6 | 13 | 12 | 9 | 14 | 4 |
| X14 | 14 | 11 | 11 | 14 | 4 | 1 | 3 | 10 | 8 | 13 | 13 |
| X15 | 11 | 3 | 1 | 15 | 10 | 5 | 8 | 3 | 6 | 3 | 3 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 12 | 3 | 10 | 4 | 5 | 10 | 13 | 9 | 12 | 13 | 5 |
| X85 | 4 | 13 | 1 | 15 | 7 | 7 | 6 | 9 | 10 | 2 | 9 |
| X86 | 8 | 14 | 13 | 6 | 10 | 7 | 1 | 4 | 4 | 9 | 9 |
| X87 | 8 | 14 | 1 | 4 | 13 | 5 | 1 | 5 | 11 | 13 | 15 |
| X88 | 12 | 4 | 13 | 15 | 2 | 9 | 2 | 3 | 10 | 6 | 15 |
| X89 | 14 | 10 | 15 | 1 | 7 | 1 | 5 | 15 | 7 | 14 | 13 |
| X90 | 3 | 6 | 12 | 8 | 4 | 7 | 5 | 5 | 12 | 15 | 12 |

**Lampiran 17.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 16 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 6 | 13 | 12 | 10 | 10 | 3 | 4 | 4 | 16 | 3 | 8 |
| X2 | 9 | 5 | 16 | 15 | 6 | 1 | 13 | 1 | 6 | 14 | 4 |
| X3 | 1 | 1 | 5 | 7 | 11 | 4 | 3 | 9 | 11 | 2 | 9 |
| X4 | 14 | 13 | 11 | 10 | 2 | 6 | 5 | 6 | 16 | 14 | 10 |
| X5 | 11 | 12 | 11 | 8 | 12 | 5 | 10 | 6 | 4 | 11 | 1 |
| X6 | 2 | 16 | 6 | 7 | 13 | 12 | 12 | 5 | 14 | 5 | 7 |
| X7 | 14 | 11 | 9 | 4 | 10 | 8 | 10 | 4 | 12 | 12 | 14 |
| X8 | 9 | 3 | 7 | 1 | 4 | 2 | 8 | 1 | 7 | 16 | 6 |
| X9 | 15 | 16 | 15 | 4 | 14 | 9 | 4 | 10 | 11 | 13 | 14 |
| X10 | 13 | 8 | 5 | 12 | 6 | 3 | 8 | 11 | 8 | 15 | 7 |
| X11 | 13 | 13 | 16 | 10 | 3 | 3 | 3 | 11 | 1 | 14 | 2 |
| X12 | 14 | 10 | 5 | 6 | 2 | 6 | 5 | 3 | 8 | 14 | 7 |
| X13 | 4 | 7 | 9 | 1 | 10 | 10 | 1 | 9 | 2 | 15 | 7 |
| X14 | 13 | 14 | 13 | 2 | 16 | 12 | 11 | 6 | 3 | 7 | 3 |
| X15 | 7 | 10 | 10 | 16 | 1 | 2 | 11 | 2 | 13 | 1 | 9 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 8 | 2 | 2 | 5 | 16 | 3 | 6 | 9 | 6 | 4 | 13 |
| X85 | 12 | 14 | 7 | 13 | 14 | 13 | 12 | 14 | 5 | 6 | 7 |
| X86 | 16 | 6 | 8 | 4 | 10 | 5 | 12 | 16 | 5 | 4 | 16 |
| X87 | 2 | 5 | 14 | 1 | 1 | 2 | 4 | 3 | 8 | 13 | 4 |
| X88 | 3 | 1 | 10 | 6 | 14 | 9 | 3 | 4 | 4 | 12 | 3 |
| X89 | 16 | 5 | 1 | 9 | 8 | 4 | 2 | 14 | 6 | 2 | 1 |
| X90 | 1 | 4 | 4 | 7 | 12 | 9 | 6 | 7 | 15 | 10 | 14 |

**Lampiran 18.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 17 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 2 | 9 | 7 | 12 | 2 | 15 | 17 | 4 | 2 | 13 | 9 |
| X2 | 7 | 1 | 12 | 1 | 5 | 7 | 8 | 3 | 12 | 14 | 2 |
| X3 | 11 | 8 | 16 | 4 | 8 | 14 | 14 | 16 | 15 | 10 | 16 |
| X4 | 15 | 6 | 17 | 11 | 12 | 14 | 14 | 9 | 15 | 12 | 10 |
| X5 | 12 | 13 | 11 | 8 | 9 | 11 | 16 | 2 | 9 | 3 | 8 |
| X6 | 16 | 14 | 16 | 3 | 2 | 5 | 12 | 4 | 15 | 15 | 12 |
| X7 | 15 | 15 | 15 | 16 | 4 | 8 | 6 | 10 | 8 | 11 | 15 |
| X8 | 4 | 12 | 2 | 14 | 13 | 13 | 11 | 11 | 12 | 17 | 5 |
| X9 | 10 | 8 | 3 | 4 | 1 | 14 | 3 | 7 | 6 | 15 | 4 |
| X10 | 4 | 11 | 15 | 15 | 8 | 10 | 6 | 13 | 6 | 1 | 16 |
| X11 | 6 | 2 | 15 | 6 | 10 | 9 | 5 | 12 | 15 | 6 | 7 |
| X12 | 14 | 4 | 6 | 9 | 17 | 10 | 9 | 9 | 17 | 13 | 11 |
| X13 | 12 | 4 | 6 | 10 | 12 | 17 | 17 | 7 | 6 | 1 | 9 |
| X14 | 6 | 5 | 3 | 5 | 16 | 9 | 16 | 15 | 17 | 2 | 4 |
| X15 | 2 | 7 | 11 | 13 | 13 | 1 | 2 | 3 | 17 | 2 | 12 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 5 | 6 | 9 | 3 | 8 | 4 | 13 | 9 | 1 | 10 | 10 |
| X85 | 3 | 2 | 16 | 14 | 14 | 2 | 10 | 12 | 7 | 13 | 11 |
| X86 | 4 | 6 | 11 | 16 | 10 | 8 | 8 | 14 | 17 | 7 | 11 |
| X87 | 1 | 1 | 17 | 6 | 1 | 12 | 1 | 8 | 2 | 2 | 8 |
| X88 | 8 | 15 | 7 | 8 | 16 | 5 | 15 | 7 | 1 | 17 | 14 |
| X89 | 7 | 2 | 2 | 6 | 15 | 10 | 1 | 15 | 9 | 2 | 6 |
| X90 | 12 | 11 | 6 | 5 | 1 | 10 | 6 | 11 | 11 | 3 | 3 |

**Lampiran 19.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 18 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 6 | 17 | 6 | 14 | 3 | 16 | 10 | 18 | 3 | 7 | 11 |
| X2 | 17 | 8 | 10 | 7 | 2 | 3 | 12 | 5 | 3 | 6 | 16 |
| X3 | 5 | 16 | 18 | 7 | 3 | 7 | 10 | 15 | 2 | 12 | 17 |
| X4 | 7 | 11 | 1 | 13 | 15 | 8 | 5 | 12 | 10 | 9 | 4 |
| X5 | 2 | 10 | 17 | 15 | 18 | 4 | 4 | 17 | 2 | 8 | 13 |
| X6 | 12 | 7 | 5 | 1 | 9 | 10 | 6 | 16 | 9 | 14 | 11 |
| X7 | 13 | 4 | 9 | 5 | 3 | 8 | 17 | 8 | 1 | 4 | 17 |
| X8 | 16 | 5 | 10 | 16 | 5 | 9 | 9 | 12 | 4 | 4 | 8 |
| X9 | 8 | 13 | 17 | 4 | 14 | 11 | 9 | 8 | 4 | 14 | 16 |
| X10 | 9 | 17 | 1 | 9 | 17 | 11 | 16 | 7 | 18 | 8 | 8 |
| X11 | 4 | 15 | 2 | 6 | 5 | 13 | 3 | 15 | 13 | 13 | 9 |
| X12 | 12 | 15 | 9 | 8 | 16 | 5 | 16 | 1 | 14 | 9 | 5 |
| X13 | 1 | 12 | 11 | 6 | 13 | 15 | 13 | 15 | 14 | 11 | 10 |
| X14 | 4 | 18 | 12 | 3 | 17 | 14 | 4 | 14 | 2 | 6 | 14 |
| X15 | 6 | 6 | 9 | 13 | 17 | 6 | 6 | 11 | 7 | 7 | 16 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 13 | 3 | 14 | 18 | 13 | 3 | 2 | 10 | 3 | 10 | 3 |
| X85 | 17 | 10 | 7 | 12 | 1 | 4 | 13 | 8 | 2 | 15 | 16 |
| X86 | 4 | 18 | 10 | 3 | 12 | 17 | 11 | 10 | 11 | 2 | 10 |
| X87 | 7 | 11 | 5 | 4 | 12 | 16 | 3 | 2 | 5 | 18 | 9 |
| X88 | 1 | 6 | 12 | 14 | 14 | 8 | 9 | 17 | 2 | 16 | 9 |
| X89 | 13 | 13 | 3 | 12 | 15 | 2 | 3 | 12 | 8 | 9 | 11 |
| X90 | 16 | 8 | 9 | 9 | 9 | 9 | 6 | 13 | 15 | 14 | 12 |

**Lampiran 20.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 19 *Subspace Threshold* 0.3 *Citotoxicity* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 13 | 1 | 10 | 5 | 9 | 15 | 8 | 18 | 14 | 14 | 13 |
| X2 | 13 | 10 | 9 | 3 | 12 | 17 | 17 | 14 | 17 | 10 | 10 |
| X3 | 19 | 14 | 17 | 7 | 13 | 18 | 8 | 5 | 6 | 3 | 6 |
| X4 | 12 | 11 | 7 | 3 | 18 | 19 | 17 | 6 | 12 | 14 | 13 |
| X5 | 7 | 13 | 5 | 11 | 7 | 16 | 18 | 3 | 18 | 18 | 2 |
| X6 | 5 | 15 | 13 | 7 | 13 | 19 | 11 | 8 | 8 | 7 | 8 |
| X7 | 15 | 7 | 5 | 1 | 7 | 9 | 14 | 2 | 17 | 14 | 9 |
| X8 | 7 | 6 | 12 | 18 | 3 | 3 | 7 | 16 | 1 | 18 | 9 |
| X9 | 14 | 11 | 9 | 19 | 5 | 12 | 3 | 11 | 17 | 12 | 18 |
| X10 | 4 | 11 | 14 | 18 | 18 | 12 | 8 | 3 | 15 | 7 | 11 |
| X11 | 14 | 2 | 2 | 1 | 10 | 8 | 4 | 18 | 14 | 1 | 4 |
| X12 | 11 | 4 | 12 | 4 | 14 | 2 | 2 | 19 | 13 | 7 | 6 |
| X13 | 3 | 14 | 17 | 2 | 5 | 4 | 19 | 7 | 19 | 6 | 13 |
| X14 | 12 | 1 | 17 | 7 | 17 | 13 | 9 | 16 | 8 | 15 | 13 |
| X15 | 4 | 1 | 16 | 15 | 10 | 1 | 19 | 11 | 4 | 18 | 10 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 16 | 19 | 11 | 19 | 3 | 16 | 1 | 11 | 4 | 16 | 7 |
| X85 | 1 | 14 | 15 | 18 | 19 | 1 | 10 | 12 | 2 | 19 | 7 |
| X86 | 18 | 2 | 14 | 1 | 15 | 8 | 6 | 7 | 17 | 15 | 12 |
| X87 | 12 | 17 | 14 | 13 | 9 | 18 | 18 | 18 | 16 | 17 | 18 |
| X88 | 13 | 16 | 7 | 18 | 5 | 4 | 10 | 9 | 8 | 18 | 16 |
| X89 | 13 | 5 | 16 | 19 | 8 | 1 | 13 | 7 | 19 | 17 | 11 |
| X90 | 15 | 15 | 13 | 6 | 5 | 15 | 2 | 10 | 3 | 12 | 4 |

**Lampiran 21.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 12 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax*

| **Compound** | ***Ensemble*** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| X1 | 12 | 1 | 5 | 9 | 10 | 11 | 3 | 4 | 6 | 10 | 9 |
| X2 | 2 | 5 | 1 | 11 | 6 | 8 | 1 | 7 | 9 | 11 | 11 |
| X3 | 8 | 7 | 4 | 4 | 2 | 9 | 10 | 10 | 10 | 4 | 6 |
| X4 | 11 | 2 | 7 | 5 | 8 | 12 | 11 | 12 | 9 | 1 | 12 |
| X5 | 5 | 8 | 9 | 9 | 4 | 5 | 11 | 5 | 6 | 6 | 11 |
| X6 | 1 | 2 | 5 | 9 | 5 | 12 | 1 | 10 | 6 | 6 | 9 |
| X7 | 1 | 4 | 4 | 6 | 7 | 2 | 1 | 10 | 11 | 1 | 6 |
| X8 | 10 | 9 | 3 | 9 | 12 | 3 | 7 | 10 | 10 | 4 | 3 |
| X9 | 11 | 10 | 2 | 11 | 3 | 8 | 8 | 5 | 3 | 8 | 5 |
| X10 | 3 | 1 | 1 | 11 | 1 | 10 | 9 | 10 | 1 | 7 | 7 |
| X11 | 12 | 8 | 3 | 2 | 7 | 8 | 5 | 1 | 1 | 11 | 12 |
| X12 | 9 | 12 | 2 | 3 | 4 | 12 | 4 | 5 | 7 | 3 | 12 |
| X13 | 3 | 3 | 6 | 2 | 3 | 3 | 3 | 12 | 8 | 6 | 5 |
| X14 | 5 | 10 | 9 | 2 | 10 | 5 | 3 | 7 | 7 | 7 | 11 |
| X15 | 6 | 7 | 9 | 2 | 5 | 2 | 6 | 5 | 12 | 9 | 2 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 5 | 10 | 8 | 2 | 9 | 9 | 2 | 6 | 7 | 11 | 5 |
| X85 | 3 | 9 | 1 | 8 | 6 | 1 | 11 | 6 | 3 | 12 | 2 |
| X86 | 9 | 11 | 5 | 2 | 9 | 10 | 5 | 3 | 5 | 3 | 9 |
| X87 | 10 | 10 | 11 | 12 | 7 | 9 | 11 | 11 | 12 | 9 | 9 |
| X88 | 8 | 10 | 12 | 10 | 1 | 3 | 6 | 12 | 3 | 3 | 8 |
| X89 | 1 | 5 | 3 | 8 | 5 | 6 | 9 | 5 | 7 | 4 | 8 |
| X90 | 3 | 12 | 2 | 11 | 6 | 3 | 4 | 6 | 3 | 8 | 12 |

**Lampiran 22.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 13 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 11 | 3 | 5 | 3 | 13 | 13 | 2 | 2 | 3 | 2 | 8 |
| X2 | 11 | 9 | 9 | 9 | 13 | 4 | 12 | 6 | 12 | 2 | 1 |
| X3 | 2 | 6 | 12 | 4 | 5 | 6 | 2 | 6 | 3 | 12 | 1 |
| X4 | 10 | 13 | 3 | 11 | 9 | 10 | 1 | 1 | 1 | 9 | 7 |
| X5 | 9 | 11 | 1 | 10 | 12 | 12 | 2 | 1 | 5 | 1 | 8 |
| X6 | 10 | 5 | 4 | 1 | 1 | 1 | 5 | 12 | 2 | 13 | 5 |
| X7 | 11 | 3 | 13 | 5 | 2 | 4 | 13 | 11 | 13 | 4 | 8 |
| X8 | 9 | 12 | 11 | 9 | 1 | 3 | 6 | 6 | 13 | 12 | 12 |
| X9 | 11 | 10 | 7 | 5 | 6 | 2 | 12 | 3 | 13 | 10 | 12 |
| X10 | 2 | 4 | 10 | 11 | 10 | 11 | 7 | 4 | 5 | 9 | 4 |
| X11 | 8 | 13 | 1 | 10 | 10 | 6 | 2 | 11 | 10 | 6 | 3 |
| X12 | 7 | 1 | 2 | 7 | 3 | 7 | 6 | 11 | 13 | 7 | 2 |
| X13 | 7 | 5 | 11 | 4 | 4 | 4 | 13 | 10 | 3 | 8 | 12 |
| X14 | 1 | 4 | 5 | 7 | 3 | 4 | 9 | 6 | 9 | 1 | 9 |
| X15 | 10 | 9 | 7 | 8 | 10 | 11 | 1 | 12 | 9 | 13 | 1 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 6 | 12 | 12 | 1 | 5 | 6 | 6 | 2 | 9 | 11 | 10 |
| X85 | 12 | 5 | 12 | 13 | 7 | 7 | 5 | 10 | 5 | 6 | 3 |
| X86 | 9 | 10 | 5 | 12 | 4 | 4 | 3 | 9 | 9 | 12 | 13 |
| X87 | 2 | 1 | 1 | 6 | 8 | 11 | 1 | 2 | 6 | 9 | 7 |
| X88 | 4 | 6 | 8 | 7 | 3 | 11 | 11 | 9 | 5 | 12 | 10 |
| X89 | 9 | 5 | 7 | 13 | 4 | 9 | 4 | 5 | 3 | 3 | 9 |
| X90 | 3 | 11 | 8 | 1 | 12 | 12 | 8 | 12 | 1 | 0:00 | 5 |

**Lampiran 23.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 14 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| X1 | 3 | 13 | 12 | 10 | 5 | 6 | 12 | 2 | 4 | 6 | 2 |
| X2 | 10 | 8 | 11 | 11 | 3 | 8 | 13 | 14 | 9 | 6 | 2 |
| X3 | 1 | 13 | 4 | 7 | 11 | 12 | 14 | 1 | 2 | 13 | 12 |
| X4 | 8 | 14 | 8 | 6 | 11 | 13 | 7 | 4 | 3 | 5 | 9 |
| X5 | 7 | 10 | 7 | 9 | 1 | 10 | 11 | 14 | 8 | 2 | 1 |
| X6 | 3 | 14 | 1 | 2 | 3 | 12 | 5 | 1 | 9 | 5 | 13 |
| X7 | 7 | 14 | 12 | 2 | 1 | 12 | 5 | 14 | 4 | 5 | 4 |
| X8 | 9 | 3 | 6 | 9 | 4 | 14 | 1 | 1 | 12 | 12 | 12 |
| X9 | 4 | 11 | 9 | 1 | 5 | 5 | 12 | 13 | 12 | 7 | 10 |
| X10 | 14 | 6 | 10 | 5 | 7 | 11 | 4 | 3 | 10 | 9 | 9 |
| X11 | 11 | 8 | 2 | 13 | 14 | 13 | 8 | 5 | 3 | 4 | 6 |
| X12 | 8 | 10 | 12 | 5 | 5 | 11 | 5 | 10 | 8 | 2 | 7 |
| X13 | 3 | 9 | 1 | 6 | 5 | 4 | 3 | 9 | 7 | 3 | 8 |
| X14 | 8 | 1 | 3 | 9 | 14 | 1 | 2 | 9 | 5 | 7 | 1 |
| X15 | 2 | 3 | 2 | 5 | 4 | 11 | 5 | 12 | 3 | 13 | 13 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 8 | 2 | 6 | 12 | 11 | 3 | 7 | 14 | 2 | 6 | 11 |
| X85 | 9 | 2 | 5 | 5 | 8 | 10 | 7 | 13 | 5 | 7 | 6 |
| X86 | 5 | 13 | 8 | 8 | 14 | 7 | 2 | 7 | 10 | 2 | 12 |
| X87 | 9 | 8 | 6 | 4 | 12 | 12 | 8 | 13 | 10 | 1 | 9 |
| X88 | 14 | 8 | 7 | 11 | 3 | 12 | 2 | 12 | 11 | 7 | 12 |
| X89 | 6 | 2 | 5 | 1 | 14 | 8 | 3 | 3 | 2 | 7 | 3 |
| X90 | 13 | 9 | 2 | 8 | 9 | 11 | 12 | 7 | 11 | 14 | 0 |

**Lampiran 24.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 15 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax*

| **Compound** | ***Ensemble*** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** |
| X1 | 15 | 15 | 3 | 8 | 10 | 3 | 11 | 11 | 3 | 1 | 4 |
| X2 | 15 | 2 | 7 | 11 | 12 | 5 | 5 | 15 | 8 | 5 | 9 |
| X3 | 3 | 14 | 5 | 2 | 9 | 9 | 10 | 6 | 5 | 13 | 2 |
| X4 | 2 | 12 | 15 | 8 | 9 | 7 | 15 | 2 | 7 | 4 | 3 |
| X5 | 4 | 2 | 10 | 13 | 10 | 1 | 9 | 9 | 6 | 12 | 8 |
| X6 | 13 | 8 | 7 | 2 | 15 | 12 | 1 | 11 | 3 | 3 | 9 |
| X7 | 2 | 11 | 8 | 12 | 8 | 15 | 12 | 1 | 3 | 9 | 4 |
| X8 | 8 | 1 | 9 | 2 | 5 | 10 | 14 | 9 | 15 | 6 | 12 |
| X9 | 15 | 14 | 13 | 8 | 6 | 1 | 9 | 5 | 13 | 10 | 12 |
| X10 | 5 | 1 | 8 | 6 | 10 | 9 | 4 | 4 | 13 | 2 | 10 |
| X11 | 9 | 3 | 12 | 13 | 8 | 2 | 11 | 3 | 7 | 1 | 3 |
| X12 | 3 | 14 | 13 | 14 | 14 | 10 | 14 | 12 | 6 | 2 | 8 |
| X13 | 10 | 8 | 12 | 14 | 12 | 9 | 6 | 11 | 7 | 3 | 7 |
| X14 | 1 | 12 | 15 | 5 | 12 | 9 | 8 | 9 | 10 | 7 | 5 |
| X15 | 11 | 11 | 14 | 10 | 14 | 5 | 6 | 14 | 9 | 2 | 3 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 14 | 5 | 6 | 5 | 4 | 14 | 3 | 8 | 4 | 5 | 2 |
| X85 | 13 | 11 | 7 | 13 | 10 | 2 | 12 | 9 | 6 | 3 | 5 |
| X86 | 7 | 8 | 3 | 2 | 6 | 15 | 14 | 4 | 2 | 1 | 10 |
| X87 | 4 | 11 | 14 | 8 | 15 | 7 | 8 | 5 | 8 | 2 | 10 |
| X88 | 6 | 11 | 2 | 14 | 15 | 2 | 9 | 2 | 15 | 4 | 11 |
| X89 | 11 | 9 | 6 | 4 | 10 | 15 | 7 | 13 | 10 | 12 | 2 |
| X90 | 5 | 6 | 4 | 15 | 8 | 15 | 13 | 14 | 9 | 9 | 11 |

**Lampiran 25.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 16 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 4 | 15 | 6 | 7 | 3 | 7 | 4 | 12 | 6 | 7 | 11 |
| X2 | 12 | 15 | 11 | 2 | 11 | 1 | 1 | 3 | 10 | 10 | 15 |
| X3 | 9 | 14 | 6 | 10 | 15 | 10 | 5 | 7 | 10 | 6 | 6 |
| X4 | 5 | 16 | 7 | 11 | 14 | 7 | 7 | 11 | 15 | 3 | 2 |
| X5 | 1 | 13 | 13 | 15 | 10 | 16 | 2 | 15 | 16 | 16 | 9 |
| X6 | 2 | 16 | 4 | 1 | 3 | 11 | 9 | 15 | 8 | 2 | 11 |
| X7 | 8 | 7 | 11 | 4 | 1 | 13 | 16 | 16 | 3 | 3 | 1 |
| X8 | 3 | 15 | 16 | 9 | 15 | 5 | 11 | 8 | 4 | 11 | 9 |
| X9 | 1 | 4 | 1 | 15 | 14 | 4 | 6 | 4 | 3 | 1 | 5 |
| X10 | 3 | 8 | 15 | 8 | 16 | 8 | 7 | 1 | 7 | 5 | 4 |
| X11 | 6 | 8 | 5 | 11 | 11 | 14 | 12 | 3 | 13 | 5 | 3 |
| X12 | 4 | 13 | 15 | 8 | 8 | 16 | 10 | 10 | 8 | 6 | 12 |
| X13 | 15 | 11 | 3 | 14 | 6 | 10 | 11 | 3 | 14 | 1 | 11 |
| X14 | 4 | 1 | 14 | 7 | 7 | 8 | 2 | 1 | 9 | 4 | 9 |
| X15 | 16 | 14 | 6 | 15 | 9 | 8 | 6 | 9 | 14 | 12 | 14 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 3 | 11 | 7 | 9 | 7 | 15 | 5 | 15 | 11 | 2 | 8 |
| X85 | 16 | 2 | 12 | 13 | 11 | 3 | 2 | 8 | 5 | 3 | 9 |
| X86 | 4 | 7 | 15 | 14 | 3 | 2 | 11 | 2 | 2 | 8 | 4 |
| X87 | 11 | 2 | 16 | 3 | 1 | 16 | 3 | 12 | 13 | 4 | 5 |
| X88 | 16 | 2 | 9 | 7 | 5 | 5 | 6 | 6 | 7 | 6 | 2 |
| X89 | 2 | 10 | 3 | 1 | 2 | 6 | 12 | 2 | 5 | 1 | 13 |
| X90 | 7 | 7 | 8 | 4 | 8 | 11 | 9 | 2 | 4 | 10 | 14 |

**Lampiran 26.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 17 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 13 | 15 | 8 | 11 | 6 | 4 | 15 | 16 | 16 | 16 | 4 |
| X2 | 2 | 12 | 6 | 3 | 7 | 8 | 5 | 7 | 8 | 3 | 1 |
| X3 | 16 | 5 | 8 | 15 | 7 | 3 | 12 | 14 | 14 | 7 | 5 |
| X4 | 10 | 4 | 3 | 15 | 5 | 1 | 5 | 5 | 10 | 14 | 7 |
| X5 | 3 | 15 | 3 | 5 | 13 | 17 | 5 | 4 | 1 | 11 | 2 |
| X6 | 15 | 7 | 13 | 13 | 15 | 3 | 13 | 8 | 14 | 13 | 9 |
| X7 | 2 | 9 | 11 | 16 | 9 | 10 | 12 | 6 | 2 | 7 | 16 |
| X8 | 5 | 3 | 9 | 17 | 3 | 3 | 8 | 8 | 14 | 5 | 11 |
| X9 | 9 | 9 | 13 | 12 | 7 | 15 | 1 | 9 | 15 | 11 | 6 |
| X10 | 14 | 1 | 1 | 6 | 12 | 13 | 1 | 17 | 15 | 11 | 7 |
| X11 | 2 | 8 | 8 | 8 | 14 | 13 | 5 | 7 | 15 | 1 | 12 |
| X12 | 13 | 9 | 8 | 16 | 10 | 11 | 15 | 2 | 3 | 10 | 10 |
| X13 | 17 | 1 | 12 | 5 | 1 | 12 | 4 | 10 | 12 | 15 | 11 |
| X14 | 8 | 5 | 16 | 13 | 8 | 13 | 6 | 3 | 9 | 2 | 2 |
| X15 | 13 | 7 | 17 | 3 | 11 | 1 | 13 | 12 | 16 | 12 | 6 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 16 | 12 | 4 | 3 | 7 | 15 | 8 | 17 | 3 | 1 | 5 |
| X85 | 14 | 3 | 4 | 17 | 15 | 8 | 6 | 17 | 4 | 15 | 2 |
| X86 | 17 | 5 | 2 | 6 | 1 | 13 | 17 | 14 | 15 | 1 | 11 |
| X87 | 14 | 15 | 2 | 11 | 11 | 1 | 9 | 3 | 9 | 3 | 3 |
| X88 | 11 | 15 | 3 | 9 | 15 | 17 | 12 | 2 | 6 | 3 | 6 |
| X89 | 2 | 17 | 13 | 8 | 9 | 1 | 9 | 1 | 7 | 9 | 12 |
| X90 | 5 | 10 | 8 | 11 | 10 | 5 | 14 | 1 | 16 | 9 | 9 |

**Lampiran 27.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 18 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 4 | 6 | 18 | 7 | 8 | 8 | 12 | 5 | 9 | 2 | 6 |
| X2 | 2 | 11 | 7 | 10 | 9 | 8 | 10 | 7 | 4 | 11 | 7 |
| X3 | 12 | 8 | 17 | 15 | 11 | 16 | 16 | 16 | 1 | 8 | 7 |
| X4 | 10 | 18 | 14 | 5 | 8 | 8 | 12 | 17 | 10 | 1 | 5 |
| X5 | 4 | 17 | 15 | 13 | 16 | 16 | 5 | 18 | 13 | 9 | 13 |
| X6 | 10 | 11 | 13 | 5 | 16 | 9 | 14 | 4 | 5 | 8 | 15 |
| X7 | 17 | 10 | 7 | 14 | 10 | 10 | 3 | 18 | 18 | 15 | 9 |
| X8 | 13 | 6 | 15 | 12 | 1 | 13 | 3 | 5 | 8 | 9 | 3 |
| X9 | 8 | 16 | 18 | 4 | 18 | 11 | 1 | 15 | 12 | 7 | 7 |
| X10 | 10 | 15 | 12 | 3 | 9 | 5 | 3 | 5 | 9 | 18 | 12 |
| X11 | 5 | 18 | 6 | 8 | 14 | 6 | 2 | 9 | 10 | 17 | 14 |
| X12 | 1 | 13 | 17 | 17 | 10 | 17 | 11 | 13 | 9 | 1 | 10 |
| X13 | 2 | 15 | 6 | 3 | 3 | 15 | 18 | 12 | 15 | 5 | 1 |
| X14 | 18 | 2 | 18 | 13 | 12 | 13 | 17 | 8 | 9 | 1 | 8 |
| X15 | 15 | 10 | 5 | 14 | 17 | 4 | 7 | 3 | 12 | 2 | 11 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 7 | 7 | 1 | 7 | 5 | 9 | 10 | 11 | 12 | 10 | 7 |
| X85 | 15 | 5 | 2 | 8 | 10 | 10 | 4 | 18 | 18 | 13 | 15 |
| X86 | 3 | 6 | 1 | 7 | 9 | 7 | 11 | 13 | 8 | 16 | 1 |
| X87 | 15 | 12 | 7 | 9 | 4 | 15 | 8 | 17 | 10 | 12 | 11 |
| X88 | 4 | 13 | 17 | 16 | 6 | 6 | 3 | 1 | 9 | 4 | 15 |
| X89 | 9 | 4 | 11 | 14 | 16 | 7 | 6 | 2 | 7 | 10 | 9 |
| X90 | 11 | 13 | 9 | 4 | 4 | 1 | 10 | 1 | 2 | 12 | 10 |

**Lampiran 28.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 19 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Softmax*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 7 | 15 | 7 | 9 | 2 | 9 | 14 | 13 | 14 | 16 | 7 |
| X2 | 14 | 5 | 5 | 12 | 9 | 1 | 13 | 8 | 1 | 16 | 10 |
| X3 | 14 | 4 | 2 | 5 | 3 | 2 | 9 | 18 | 18 | 14 | 15 |
| X4 | 1 | 19 | 10 | 13 | 16 | 15 | 1 | 6 | 3 | 14 | 5 |
| X5 | 1 | 17 | 1 | 8 | 8 | 11 | 2 | 18 | 10 | 13 | 13 |
| X6 | 13 | 7 | 11 | 13 | 1 | 15 | 13 | 13 | 11 | 1 | 5 |
| X7 | 3 | 1 | 19 | 5 | 18 | 10 | 1 | 19 | 10 | 17 | 14 |
| X8 | 18 | 11 | 13 | 1 | 14 | 18 | 3 | 1 | 13 | 5 | 12 |
| X9 | 14 | 4 | 15 | 19 | 1 | 6 | 2 | 12 | 12 | 10 | 4 |
| X10 | 13 | 9 | 16 | 14 | 15 | 6 | 15 | 3 | 14 | 17 | 3 |
| X11 | 13 | 14 | 19 | 3 | 18 | 14 | 16 | 9 | 6 | 5 | 8 |
| X12 | 5 | 9 | 7 | 14 | 4 | 1 | 6 | 16 | 5 | 1 | 17 |
| X13 | 19 | 4 | 5 | 5 | 5 | 16 | 10 | 8 | 1 | 11 | 3 |
| X14 | 6 | 12 | 8 | 16 | 16 | 4 | 5 | 16 | 11 | 13 | 13 |
| X15 | 3 | 16 | 6 | 5 | 2 | 18 | 14 | 16 | 14 | 8 | 14 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 14 | 8 | 6 | 11 | 5 | 1 | 5 | 5 | 11 | 4 | 7 |
| X85 | 19 | 16 | 8 | 19 | 13 | 12 | 12 | 12 | 17 | 14 | 8 |
| X86 | 13 | 11 | 10 | 15 | 2 | 5 | 5 | 12 | 1 | 7 | 7 |
| X87 | 17 | 12 | 14 | 18 | 8 | 10 | 7 | 6 | 9 | 18 | 9 |
| X88 | 5 | 6 | 16 | 9 | 8 | 10 | 15 | 9 | 14 | 14 | 16 |
| X89 | 18 | 13 | 9 | 16 | 11 | 3 | 18 | 12 | 15 | 17 | 14 |
| X90 | 10 | 5 | 8 | 2 | 14 | 8 | 12 | 9 | 13 | 19 | 4 |

**Lampiran 29.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 12 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 3 | 4 | 6 | 10 | 9 | 2 | 3 | 6 | 1 | 6 | 13 |
| X2 | 1 | 5 | 3 | 6 | 5 | 2 | 7 | 2 | 7 | 12 | 8 |
| X3 | 4 | 6 | 9 | 7 | 3 | 2 | 10 | 11 | 4 | 3 | 18 |
| X4 | 12 | 4 | 6 | 4 | 4 | 10 | 1 | 4 | 5 | 12 | 6 |
| X5 | 9 | 8 | 11 | 5 | 9 | 5 | 11 | 6 | 8 | 10 | 18 |
| X6 | 3 | 5 | 2 | 3 | 7 | 4 | 10 | 3 | 1 | 6 | 13 |
| X7 | 6 | 7 | 3 | 11 | 12 | 5 | 7 | 2 | 6 | 7 | 19 |
| X8 | 6 | 11 | 5 | 9 | 11 | 1 | 10 | 7 | 12 | 9 | 1 |
| X9 | 1 | 1 | 4 | 5 | 3 | 11 | 12 | 12 | 9 | 9 | 12 |
| X10 | 2 | 9 | 9 | 11 | 7 | 9 | 3 | 5 | 2 | 4 | 3 |
| X11 | 7 | 3 | 8 | 4 | 7 | 1 | 3 | 10 | 4 | 3 | 9 |
| X12 | 11 | 5 | 1 | 6 | 12 | 6 | 3 | 9 | 3 | 4 | 16 |
| X13 | 9 | 4 | 4 | 5 | 5 | 9 | 12 | 4 | 7 | 8 | 8 |
| X14 | 2 | 12 | 9 | 4 | 9 | 11 | 12 | 12 | 2 | 9 | 16 |
| X15 | 2 | 2 | 6 | 12 | 1 | 11 | 6 | 3 | 1 | 12 | 16 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 5 | 9 | 1 | 7 | 11 | 10 | 9 | 7 | 11 | 4 | 5 |
| X85 | 10 | 5 | 3 | 10 | 3 | 9 | 11 | 11 | 7 | 12 | 12 |
| X86 | 5 | 4 | 4 | 9 | 11 | 6 | 4 | 9 | 6 | 3 | 12 |
| X87 | 11 | 9 | 5 | 11 | 4 | 10 | 8 | 3 | 10 | 2 | 6 |
| X88 | 8 | 12 | 1 | 1 | 8 | 8 | 8 | 8 | 7 | 11 | 9 |
| X89 | 9 | 8 | 6 | 2 | 4 | 6 | 8 | 6 | 11 | 6 | 12 |
| X90 | 7 | 2 | 7 | 1 | 12 | 4 | 4 | 11 | 3 | 5 | 9 |

**Lampiran 30.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 13 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 11 | 3 | 5 | 3 | 13 | 13 | 2 | 2 | 3 | 8 | 3 |
| X2 | 11 | 9 | 9 | 9 | 13 | 4 | 12 | 6 | 12 | 1 | 7 |
| X3 | 2 | 6 | 12 | 4 | 5 | 6 | 2 | 6 | 3 | 1 | 10 |
| X4 | 10 | 13 | 3 | 11 | 9 | 10 | 1 | 1 | 1 | 7 | 1 |
| X5 | 9 | 11 | 1 | 10 | 12 | 12 | 2 | 1 | 5 | 8 | 11 |
| X6 | 10 | 5 | 4 | 1 | 1 | 1 | 5 | 12 | 2 | 5 | 10 |
| X7 | 11 | 3 | 13 | 5 | 2 | 4 | 13 | 11 | 13 | 8 | 7 |
| X8 | 9 | 12 | 11 | 9 | 1 | 3 | 6 | 6 | 13 | 12 | 10 |
| X9 | 11 | 10 | 7 | 5 | 6 | 2 | 12 | 3 | 13 | 12 | 12 |
| X10 | 2 | 4 | 10 | 11 | 10 | 11 | 7 | 4 | 5 | 4 | 3 |
| X11 | 8 | 13 | 1 | 10 | 10 | 6 | 2 | 11 | 10 | 3 | 3 |
| X12 | 7 | 1 | 2 | 7 | 3 | 7 | 6 | 11 | 13 | 2 | 3 |
| X13 | 7 | 5 | 11 | 4 | 4 | 4 | 13 | 10 | 3 | 12 | 12 |
| X14 | 1 | 4 | 5 | 7 | 3 | 4 | 9 | 6 | 9 | 9 | 12 |
| X15 | 10 | 9 | 7 | 8 | 10 | 11 | 1 | 12 | 9 | 1 | 6 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 6 | 12 | 12 | 1 | 5 | 6 | 6 | 2 | 9 | 10 | 9 |
| X85 | 12 | 5 | 12 | 13 | 7 | 7 | 5 | 10 | 5 | 3 | 11 |
| X86 | 9 | 10 | 5 | 12 | 4 | 4 | 3 | 9 | 9 | 13 | 4 |
| X87 | 2 | 1 | 1 | 6 | 8 | 11 | 1 | 2 | 6 | 7 | 8 |
| X88 | 4 | 6 | 8 | 7 | 3 | 11 | 11 | 9 | 5 | 10 | 8 |
| X89 | 9 | 5 | 7 | 13 | 4 | 9 | 4 | 5 | 3 | 9 | 8 |
| X90 | 3 | 11 | 8 | 1 | 12 | 12 | 8 | 12 | 1 | 5 | 4 |

**Lampiran 31.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 14 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 12 | 8 | 12 | 8 | 10 | 12 | 4 | 2 | 3 | 5 | 9 |
| X2 | 14 | 11 | 6 | 3 | 8 | 7 | 14 | 6 | 7 | 2 | 5 |
| X3 | 14 | 4 | 6 | 3 | 6 | 10 | 1 | 13 | 8 | 14 | 3 |
| X4 | 12 | 13 | 7 | 3 | 13 | 12 | 7 | 10 | 14 | 2 | 4 |
| X5 | 5 | 3 | 6 | 6 | 9 | 10 | 1 | 3 | 4 | 6 | 9 |
| X6 | 7 | 1 | 8 | 14 | 6 | 7 | 13 | 5 | 4 | 1 | 7 |
| X7 | 4 | 3 | 3 | 8 | 4 | 10 | 9 | 3 | 13 | 11 | 12 |
| X8 | 6 | 11 | 5 | 2 | 10 | 8 | 1 | 8 | 7 | 3 | 11 |
| X9 | 4 | 14 | 10 | 14 | 4 | 14 | 14 | 13 | 4 | 10 | 3 |
| X10 | 7 | 4 | 8 | 7 | 9 | 10 | 5 | 10 | 9 | 5 | 7 |
| X11 | 3 | 2 | 5 | 9 | 11 | 13 | 6 | 3 | 14 | 5 | 7 |
| X12 | 9 | 12 | 2 | 1 | 13 | 7 | 2 | 11 | 9 | 8 | 12 |
| X13 | 12 | 14 | 11 | 8 | 11 | 4 | 9 | 2 | 1 | 8 | 5 |
| X14 | 3 | 7 | 2 | 4 | 7 | 5 | 13 | 13 | 5 | 8 | 9 |
| X15 | 1 | 5 | 11 | 2 | 8 | 10 | 14 | 10 | 5 | 14 | 1 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 8 | 6 | 12 | 6 | 14 | 14 | 4 | 12 | 2 | 10 | 11 |
| X85 | 14 | 1 | 7 | 9 | 5 | 1 | 13 | 4 | 8 | 5 | 3 |
| X86 | 10 | 10 | 12 | 7 | 3 | 10 | 14 | 12 | 11 | 4 | 11 |
| X87 | 7 | 13 | 14 | 2 | 7 | 12 | 2 | 5 | 11 | 4 | 4 |
| X88 | 14 | 10 | 10 | 2 | 9 | 8 | 12 | 9 | 9 | 5 | 8 |
| X89 | 1 | 6 | 9 | 1 | 12 | 3 | 7 | 2 | 7 | 14 | 4 |
| X90 | 11 | 12 | 7 | 7 | 1 | 11 | 9 | 1 | 13 | 5 | 12 |

**Lampiran 32.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 15 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 15 | 10 | 7 | 8 | 8 | 10 | 4 | 9 | 12 | 13 | 10 |
| X2 | 2 | 11 | 13 | 8 | 13 | 4 | 5 | 15 | 6 | 12 | 8 |
| X3 | 0:00 | 2 | 12 | 13 | 4 | 11 | 15 | 8 | 10 | 10 | 6 |
| X4 | 12 | 7 | 10 | 1 | 1 | 11 | 15 | 1 | 7 | 1 | 13 |
| X5 | 3 | 11 | 10 | 4 | 7 | 11 | 7 | 2 | 6 | 8 | 9 |
| X6 | 8 | 2 | 8 | 5 | 11 | 9 | 11 | 9 | 9 | 15 | 6 |
| X7 | 3 | 7 | 7 | 9 | 1 | 3 | 6 | 4 | 5 | 6 | 4 |
| X8 | 3 | 7 | 11 | 12 | 9 | 3 | 2 | 7 | 12 | 10 | 10 |
| X9 | 10 | 6 | 12 | 11 | 14 | 12 | 6 | 1 | 6 | 10 | 4 |
| X10 | 8 | 11 | 4 | 13 | 10 | 14 | 14 | 4 | 9 | 8 | 9 |
| X11 | 11 | 3 | 15 | 1 | 8 | 13 | 3 | 4 | 5 | 2 | 11 |
| X12 | 5 | 1 | 9 | 14 | 4 | 8 | 7 | 5 | 2 | 10 | 13 |
| X13 | 14 | 7 | 3 | 5 | 2 | 10 | 12 | 11 | 13 | 7 | 11 |
| X14 | 5 | 15 | 3 | 15 | 10 | 1 | 4 | 12 | 5 | 1 | 7 |
| X15 | 7 | 5 | 11 | 7 | 15 | 14 | 7 | 8 | 5 | 10 | 8 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 11 | 8 | 11 | 1 | 13 | 5 | 10 | 1 | 4 | 12 | 14 |
| X85 | 7 | 11 | 7 | 4 | 10 | 8 | 14 | 9 | 12 | 12 | 5 |
| X86 | 1 | 8 | 2 | 3 | 8 | 9 | 9 | 2 | 12 | 6 | 3 |
| X87 | 9 | 10 | 2 | 2 | 4 | 3 | 3 | 9 | 2 | 9 | 7 |
| X88 | 4 | 11 | 14 | 8 | 1 | 13 | 3 | 7 | 10 | 13 | 9 |
| X89 | 7 | 2 | 3 | 12 | 10 | 9 | 7 | 4 | 5 | 13 | 12 |
| X90 | 13 | 14 | 11 | 12 | 11 | 14 | 8 | 12 | 2 | 3 | 1 |

**Lampiran 33.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 16 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 5 | 1 | 8 | 4 | 9 | 14 | 16 | 10 | 2 | 10 | 4 |
| X2 | 7 | 4 | 8 | 8 | 2 | 4 | 7 | 5 | 12 | 12 | 5 |
| X3 | 9 | 9 | 13 | 11 | 16 | 2 | 7 | 7 | 15 | 15 | 15 |
| X4 | 3 | 2 | 15 | 4 | 8 | 10 | 8 | 16 | 2 | 6 | 15 |
| X5 | 4 | 16 | 15 | 7 | 9 | 14 | 10 | 3 | 11 | 14 | 7 |
| X6 | 4 | 1 | 6 | 16 | 9 | 3 | 7 | 13 | 15 | 10 | 11 |
| X7 | 14 | 13 | 16 | 2 | 6 | 12 | 5 | 14 | 10 | 7 | 6 |
| X8 | 4 | 4 | 9 | 16 | 10 | 16 | 6 | 3 | 10 | 10 | 2 |
| X9 | 10 | 1 | 6 | 13 | 14 | 11 | 11 | 6 | 14 | 11 | 6 |
| X10 | 16 | 3 | 5 | 12 | 15 | 3 | 14 | 13 | 9 | 5 | 14 |
| X11 | 10 | 3 | 11 | 8 | 6 | 9 | 1 | 11 | 14 | 5 | 3 |
| X12 | 4 | 10 | 5 | 15 | 3 | 1 | 10 | 8 | 2 | 14 | 7 |
| X13 | 2 | 7 | 12 | 5 | 9 | 4 | 7 | 13 | 6 | 1 | 12 |
| X14 | 11 | 7 | 4 | 9 | 14 | 4 | 6 | 1 | 6 | 12 | 4 |
| X15 | 3 | 11 | 12 | 15 | 2 | 6 | 3 | 2 | 9 | 2 | 7 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 10 | 6 | 6 | 3 | 4 | 1 | 10 | 11 | 3 | 15 | 10 |
| X85 | 13 | 13 | 2 | 9 | 13 | 14 | 13 | 10 | 14 | 7 | 14 |
| X86 | 1 | 5 | 2 | 7 | 12 | 4 | 13 | 11 | 13 | 13 | 9 |
| X87 | 9 | 1 | 7 | 4 | 12 | 16 | 6 | 12 | 4 | 8 | 3 |
| X88 | 14 | 15 | 5 | 6 | 11 | 12 | 3 | 8 | 1 | 12 | 3 |
| X89 | 13 | 12 | 15 | 7 | 7 | 2 | 16 | 8 | 5 | 4 | 7 |
| X90 | 9 | 14 | 5 | 15 | 3 | 11 | 14 | 1 | 15 | 6 | 8 |

**Lampiran 34.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 17 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 4 | 4 | 15 | 6 | 14 | 4 | 2 | 17 | 15 | 12 | 9 |
| X2 | 16 | 15 | 7 | 3 | 17 | 14 | 12 | 15 | 1 | 5 | 2 |
| X3 | 17 | 13 | 11 | 12 | 4 | 15 | 13 | 9 | 10 | 16 | 16 |
| X4 | 8 | 16 | 5 | 1 | 14 | 15 | 13 | 15 | 8 | 2 | 8 |
| X5 | 13 | 14 | 3 | 16 | 16 | 12 | 7 | 3 | 13 | 1 | 9 |
| X6 | 10 | 5 | 6 | 16 | 12 | 14 | 6 | 9 | 7 | 9 | 9 |
| X7 | 8 | 16 | 14 | 13 | 16 | 8 | 15 | 2 | 10 | 15 | 6 |
| X8 | 1 | 15 | 8 | 10 | 14 | 2 | 1 | 16 | 13 | 16 | 10 |
| X9 | 7 | 10 | 1 | 3 | 3 | 17 | 13 | 14 | 15 | 8 | 14 |
| X10 | 9 | 16 | 1 | 17 | 12 | 7 | 12 | 5 | 3 | 1 | 15 |
| X11 | 2 | 5 | 11 | 6 | 8 | 7 | 9 | 8 | 6 | 4 | 6 |
| X12 | 16 | 8 | 11 | 12 | 16 | 5 | 2 | 13 | 10 | 7 | 3 |
| X13 | 1 | 3 | 8 | 9 | 11 | 4 | 14 | 9 | 3 | 12 | 9 |
| X14 | 16 | 10 | 14 | 10 | 1 | 10 | 1 | 14 | 4 | 11 | 14 |
| X15 | 13 | 9 | 17 | 16 | 11 | 7 | 10 | 10 | 10 | 8 | 2 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 12 | 11 | 6 | 14 | 16 | 6 | 13 | 10 | 12 | 7 | 4 |
| X85 | 13 | 10 | 5 | 11 | 7 | 13 | 3 | 2 | 1 | 13 | 13 |
| X86 | 1 | 7 | 11 | 5 | 5 | 10 | 5 | 12 | 5 | 3 | 12 |
| X87 | 3 | 9 | 4 | 17 | 2 | 7 | 17 | 7 | 8 | 3 | 12 |
| X88 | 6 | 17 | 4 | 14 | 16 | 10 | 10 | 5 | 15 | 17 | 11 |
| X89 | 12 | 15 | 9 | 8 | 5 | 16 | 17 | 11 | 12 | 9 | 7 |
| X90 | 3 | 11 | 1 | 2 | 2 | 16 | 10 | 6 | 2 | 17 | 3 |

**Lampiran 35.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 18 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 5 | 15 | 7 | 16 | 14 | 4 | 3 | 18 | 10 | 9 | 17 |
| X2 | 9 | 10 | 7 | 5 | 8 | 12 | 13 | 4 | 7 | 14 | 15 |
| X3 | 15 | 8 | 11 | 15 | 9 | 5 | 17 | 12 | 5 | 16 | 9 |
| X4 | 7 | 7 | 11 | 8 | 2 | 18 | 9 | 7 | 17 | 11 | 15 |
| X5 | 12 | 11 | 18 | 15 | 5 | 3 | 5 | 15 | 4 | 8 | 3 |
| X6 | 10 | 2 | 13 | 15 | 9 | 14 | 8 | 7 | 2 | 7 | 9 |
| X7 | 2 | 11 | 17 | 4 | 17 | 5 | 10 | 1 | 8 | 1 | 2 |
| X8 | 16 | 17 | 4 | 5 | 4 | 2 | 9 | 2 | 12 | 18 | 16 |
| X9 | 9 | 18 | 4 | 3 | 11 | 2 | 6 | 17 | 2 | 13 | 14 |
| X10 | 11 | 17 | 12 | 2 | 15 | 9 | 2 | 3 | 10 | 10 | 5 |
| X11 | 7 | 7 | 2 | 3 | 4 | 16 | 6 | 11 | 8 | 1 | 8 |
| X12 | 15 | 9 | 15 | 16 | 1 | 12 | 10 | 17 | 5 | 7 | 13 |
| X13 | 11 | 2 | 11 | 16 | 11 | 16 | 16 | 5 | 1 | 3 | 9 |
| X14 | 11 | 12 | 9 | 7 | 15 | 18 | 13 | 4 | 14 | 15 | 14 |
| X15 | 8 | 7 | 10 | 18 | 4 | 14 | 3 | 9 | 1 | 13 | 10 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 14 | 14 | 6 | 17 | 18 | 11 | 12 | 17 | 6 | 15 | 10 |
| X85 | 2 | 18 | 3 | 10 | 14 | 1 | 13 | 1 | 10 | 9 | 2 |
| X86 | 5 | 4 | 18 | 11 | 5 | 5 | 18 | 17 | 7 | 11 | 12 |
| X87 | 8 | 1 | 3 | 1 | 1 | 10 | 15 | 12 | 11 | 1 | 7 |
| X88 | 9 | 8 | 15 | 1 | 17 | 7 | 12 | 14 | 12 | 9 | 5 |
| X89 | 3 | 11 | 18 | 13 | 15 | 4 | 7 | 12 | 10 | 1 | 11 |
| X90 | 6 | 16 | 10 | 18 | 7 | 5 | 16 | 6 | 11 | 18 | 6 |

**Lampiran 36.** Alokasi Variabel Prediktor Pada *ANN Ensemble* 19 *Subspace Threshold* 0.625 *Radioprotective* dengan Aktivasi *Sigmoid*

| Compound | *Ensemble* | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
| X1 | 2 | 2 | 14 | 14 | 11 | 9 | 2 | 3 | 2 | 5 | 10 |
| X2 | 11 | 3 | 8 | 15 | 5 | 14 | 1 | 5 | 8 | 4 | 7 |
| X3 | 10 | 14 | 10 | 18 | 3 | 2 | 19 | 9 | 16 | 11 | 5 |
| X4 | 18 | 17 | 15 | 8 | 12 | 1 | 12 | 14 | 19 | 8 | 17 |
| X5 | 7 | 13 | 19 | 7 | 13 | 15 | 8 | 14 | 16 | 13 | 4 |
| X6 | 17 | 4 | 5 | 16 | 5 | 18 | 13 | 14 | 8 | 18 | 2 |
| X7 | 13 | 10 | 7 | 4 | 19 | 18 | 2 | 17 | 12 | 11 | 8 |
| X8 | 5 | 1 | 6 | 6 | 18 | 10 | 4 | 4 | 8 | 17 | 12 |
| X9 | 13 | 2 | 18 | 18 | 8 | 5 | 4 | 19 | 11 | 7 | 2 |
| X10 | 4 | 5 | 1 | 18 | 3 | 19 | 4 | 11 | 15 | 4 | 10 |
| X11 | 12 | 10 | 14 | 5 | 7 | 2 | 17 | 2 | 3 | 10 | 8 |
| X12 | 15 | 12 | 19 | 13 | 5 | 14 | 12 | 9 | 12 | 9 | 5 |
| X13 | 7 | 17 | 6 | 12 | 8 | 13 | 14 | 17 | 5 | 19 | 1 |
| X14 | 9 | 16 | 13 | 13 | 3 | 17 | 12 | 11 | 17 | 2 | 14 |
| X15 | 1 | 18 | 5 | 6 | 12 | 16 | 16 | 2 | 9 | 4 | 1 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| X84 | 1 | 11 | 17 | 12 | 10 | 17 | 12 | 3 | 19 | 7 | 6 |
| X85 | 1 | 4 | 13 | 15 | 17 | 11 | 11 | 10 | 16 | 17 | 10 |
| X86 | 19 | 11 | 7 | 4 | 10 | 3 | 7 | 8 | 15 | 19 | 7 |
| X87 | 5 | 4 | 9 | 19 | 14 | 1 | 9 | 11 | 19 | 3 | 11 |
| X88 | 4 | 9 | 4 | 3 | 10 | 16 | 16 | 12 | 11 | 9 | 12 |
| X89 | 14 | 16 | 4 | 16 | 16 | 5 | 9 | 16 | 15 | 3 | 10 |
| X90 | 11 | 8 | 19 | 6 | 16 | 11 | 10 | 16 | 10 | 9 | 11 |

**Lampiran 37.** Surat Keterangan Pengambilan Data

**SURAT PERNYATAAN**

Saya yang bertanda tangan di bawah ini. mahasiswa Departemen Statistika FMKSD ITS.

Nama : Dian Rizky Maulina
NRP : 06211540000091

menyatakan bahwa data yang digunakan dalam Tugas Akhir ini merupakan data sekunder yang diambil dari penelitian yaitu :

| | |
|---|---|
| Sumber | : Data penelitian Motohashi *et al*. (2018) dengan judul "*Regression Models and Ranking Method for P53 Inhibiotr Candidates Using Machine Learning*" |
| Keterangan | : Data kadar toksisitas dan radio proteksi dengan 20 observasi (senyawa) serta 217 variabel prediktor yang merupakan penyusun senyawa. |

Surat pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Surabaya. Juli 2019

Mengetahui,
Pembimbing Tugas Akhir

Dr. rer. pol. Heri Kuswanto, S.Si., M.Si
NIP. 19820326 200312 1 004

Mahasiswa

Dian Rizky Maulina
NRP. 06211540000091

# BIODATA PENULIS

Penulis dengan nama lengkap Dian Rizky Maulina dilahirkan di Kediri pada 15 Desember 1995. Penulis menempuh pendidikan formal di SDN Nambakan, MTsN II Kediri, dan SMAN 2 Kediri. Penulis diterima sebagai Mahasiswa Departemen Statistika ITS melalui jalur SBMPTN pada tahun 2015. Penulis meruapakan putri dari pasangan Bapak Slamet S.Pd dan Ibu Dra. Tri Wahyuni ini aktif dalam berbagai organisasi seperti Anggota KPU HIMASTA-ITS sejak 2016 hingga 2018, *Staff* Keilmiahan HIMASTA-ITS 2016/2017, *Staff* Hubungan Luar BEM-ITS 2017/2018, PIC Region Kediri untuk Station 2017, dan Laison Officer peserta luar negeri DAC 2017. Penulis aktif dalam kegiatan internasionalisasi seperti menjadi kandidat *ITS Goes Global to Thailand* pada April 2018 dan kandidat dalam *Kumamoto University International Summer Program 2019*. Selama menjalani masa perkuliahan, penulis pernah mengikuti beberapa *project* sebagai *surveyor* dan juga sebagai pengajar pada beberapa *private course*. Penulis juga berkesempatan untuk melakukan *internship program* pada PT. Astra Daihatsu Motor di Jakarta pada bulan Juli-Agustus 2018. Pada saat internship, penulis berkesempatan mengerjakan beberapa *project* yaitu *project* desain Daihatsu *we care, project* data *unit in operation* (UIO), dan *project SRO Training*. Selain itu, penulis juga berkesempatan berkontribusi dalam kegiatan GIIAS 2018 yang bertempat di ICE BSD, Tangerang. Apabila pembaca ingin memberi kritik dan saran serta diskusi lebih lanjut mengenai Tugas Akhir ini, dapat menghubungi penulis melalui email dianrmaulina@gmail.com atau nomor telepon 081234986785.