TESIS - BM185407

REKOMENDASI KEBIJAKAN STANDARISASI MENU HAK AKSES DENGAN PENDEKATAN PENGKLASTERAN PERAN PADA SISTEM PERBANKAN XYZ

YUDHISTIRO TRAH KUSUMONEGORO
09211550053006

Dosen Pembimbing:
Dr. Eng. Febriliyan Samopa, SKom, MKom

Departemen Manajemen Teknologi
Fakultas Bisnis Dan Manajemen Teknologi
Institut Teknologi Sepuluh Nopember
2019

LEMBAR PENGESAHAN TESIS

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar

Magister Manajemen Teknologi (M.MT)

di

Institut Teknologi Sepuluh Nopember

Oleh:

Yudhistiro Trah Kusumonegoro

NRP: 09211550053006

Tanggal Ujian: 11 Juli 2019

Periode Wisuda: September 2019

Disetujui oleh:

Pembimbing:

1. **Dr.Eng. Febriliyan Samopa, S.Kom M.Kom**
   NIP: 197302191998021001

.....

Penguji:

1. **Prof. Dr. Ir.Joko Lianto Buliali M.Sc.**
   NIP: 196707271992031002

.....

2. **Faizal Mahananto, S.Kom M.Eng Ph.D.**
   NIPH: 5200201301010

.....

Kepala Departemen Manajemen Teknologi

Fakultas Bisnis dan Manajemen Teknologi

Prof. Ir. I Nyoman Pujawan, M.Eng, Ph.D, CSCP

NIP: 196912311994121076

REKOMENDASI KEBIJAKAN STANDARISASI MENU HAK AKSES DENGAN PENDEKATAN PENGKLASTERAN PERAN PADA SISTEM PERBANKAN XYZ

Nama Mahasiswa : Yudhistiro Trah Kusumonegoro
NRP : 09211550053006
Pembimbing : Dr. Eng. Febriliyan Samopa, S.Kom, M.Kom

ABSTRAK

Operasional layanan bank didukung sistem informasi perbankan disebut *Core Banking System* (CBS) dengan salah satu metode pengamanan adalah melalui pengaturan hak akses menu dari akun pengguna. Adanya permintaan operasional kantor cabang untuk perubahan hak akses tanpa suatu rujukan baku menimbulkan risiko hak akses yang tidak seharusnya dimiliki. Kasus Bank XYZ yang menjadi fokus dalam penelitian ini memiliki kantor cabang dengan beberapa tingkatan kegiatan operasional. Kesulitan pengelompokkan dikarenakan pengguna memiliki peran utama sesuai kodifikasi penamaan serta sejumlah menu untuk peran lain yang dirangkap. Tahapan dalam penelitian adalah persiapan data agar hak akses pengguna tercatat di CBS dapat diproses, analisis pengklasteran untuk mendapatkan kelompok yang optimal serta evaluasi kebenaran hasil sebagai tingkat persetujuan dari pihak manajemen Bank XYZ.

Dalam penelitian ini dilakukan dua pendekatan pengklasteran, pertama adalah pengklasteran menggunakan Fuzzy C-Means (FCM) dengan data yang telah melalui transformasi t-SNE, untuk selanjutnya dikelompokkan lebih lanjut menggunakan DBSCAN. Pendekatan kedua menggunakan Fuzzy C-Means (FCM) dengan data yang tidak melalui transformasi t-SNE, untuk selanjutnya dikelompokkan lebih lanjut menggunakan FCM kembali, sehingga dapat disebut pengklasteran FCM bertingkat. Hasil yang diperoleh menunjukkan bahwa rekomendasi standarisasi hak akses dapat dilakukan dengan kedua pendekatan pengklasteran. Pengklasteran dengan transformasi t-SNE pada fitur menu dan status pengguna memberikan hasil visual untuk membantu validasi. Sedangkan pendekatan pengklasteran FCM bertingkat memberikan kesalahan pelabelan terendah. Kesalahan dalam pengklasteran menggunakan pendekatan pertama terjadi khususnya pada pengguna dengan peran ganda karena status pengguna menjadi fitur dalam pengklasteran. Sedangkan Pengklasteran berdasarkan menu saja membuat beberapa pengguna dengan peran ganda menjadi data outlier dengan nilai keanggotaan rendah.

Hasil rekomendasi hak akses selanjutnya dapat digunakan sebagai rujukan untuk menyusun kebijakan standarisasi hak akses menu pengguna di Bank XYZ, dengan perhatian khusus pada hak ases terbatas. Adanya hak akses yang tidak seharusnya dimiliki perlu ditindaklanjuti untuk mencegah penyalahgunaan

Kata kunci—penggalian peran, role-based access control, modifikasi pengklasteran

POLICY RECOMMENDATION IN STANDARDIZING ACCESS RIGHT MENU USING ROLE CLUSTERING APPROACH IN XYZ BANKING SYSTEM

By : Yudhistiro Trah Kusumonegoro
Student Identity Number : 09211550053006
Supervisor : Dr. Eng. Febriliyan Samopa, S.Kom, M.Kom

ABSTRACT

Banking operational services are supported by a banking information system called *Core Banking System* (CBS). One of its security measure is by controlling access right assigned to user account. Due to frequent request from branches without accompanied by standard references, risk of improperly granting of access right occurs. In the Case of Bank XYZ, the subject in this research, the bank has dozens of branches with multi-tiered operational activities. While users has their account already codified into specific pattern according to one's role, it is still difficult in grouping user accounts into their proper roles as current access rights are greatly vary and situation exists where a particular user has multiple roles. It is expected that this research will yield effective access rights group, categorized as mandatory, optional and restricted, in order to identify user account with multiple roles. Steps in this research are: preparing data in order to process user access rights recorded in CBS, clustering analysis to obtain optimum grouping of access rights, and evaluation of clustering results to be proposed to Bank XYZ management.

There are two approaches used in this research, first is Fuzzy C-Means (FCM) clustering with data that are transformed using t-SNE, to be grouped further using DBSCAN. Second approach is using Fuzzy C-Means (FCM) clustering without transforming the data with t-SNE, to be grouped further with FCM again, hence this approach can be called multi-staged FCM. Clustering results demonstrates that access right reccommendation can be produced using both aproaches. Clustering with t-SNE transformation on menu features and user classes yields visual cues to aid validation. On the other hand, multi-staged FCM approach produces the least labelling error. Clustering error in the first approach occurs especially to users with multiple roles, due to the user class is a feature within clustering. Clustering with only menus, however, makes some users with multiple roles becomes outliers with low membership value.

The access rights recommendation results then can be utilized further as a referrence in developing standardized user access right policy in Bank XYZ, with special care on restricted access rights. The existence of unsanctioned access right must be investigated further to avoid misuse.

Keywords—role mining, role-based access control, modified clustering

KATA PENGANTAR

Puji syukur kepada Allah SWT atas rahmat dan anugerah-Nya yang diberikan kepada penulis sehingga tesis ini dapat diselesaikan. Tesis ini disusun sebagai syarat untuk menyelesaikan progam Pasca Sarjana (S2) Magister Manajemen Teknologi dengan Bidang Keahlian Manajemen Teknologi Informasi di ITS Surabaya.

Penyusunan tesis ini tidak terlepas dari dukungan berbagai pihak baik secara moral maupun material. Penulis menucapkan terima kasih yang sedalam-dalamnya kepada :

1. Bapak Dr. Eng. Febriliyan Samopa, S.Kom, M.Kom. selaku dosen pembimbing yang telah meluangkan waktu, tenaga dan pikiran untuk memberikan bimbingan, petunjuk dan pengarahan dalam penyelesaian tesis ini.
2. Seluruh dosen pengajar yang telah memberikan bimbingan dan pengajaran serta seluruh pimpinan dan karyawan MMT ITS yang telah banyak membantu dalam berbagai hal selama masa perkuliahan.
3. Seluruh keluarga, khususnya istri tercinta, Diana Purwitasari, yang tidak pernah berhenti memberikan dukungan dan doa.
4. Rekan-rekan kuliah MMT ITS 2015 yang telah memberikan dukungannnya sehingga dapat menyelesaikan penulisan ini.

Surabaya, Agustus 2019

(Penulis)

DAFTAR ISI

DAFTAR GAMBAR

DAFTAR TABEL

BAB 1 PENDAHULUAN

1.1 Latar Belakang

Salah satu prinsip pengamanan informasi adalah melalui pengendalian akses atas informasi tersebut berdasarkan *need to know* dan *least privilege*. Penerapan prinsip tersebut berarti memastikan bahwa akses hanya diberikan pada pihak yang membutuhkan dan tingkat akses disesuaikan dengan kebutuhan. Operasional layanan bank telah didukung dengan sistem informasi perbankan yang disebut sebagai *Core Banking System* (CBS). Sistem tersebut bersifat kritis karena mendukung layanan inti perbankan seperti tabungan, pinjaman, hingga pembukuan transaksi. Oleh karena itu pengamanan akses CBS harus dilakukan dengan seksama. Salah satu metode pengendalian akses CBS adalah melalui pengaturan akses menu. Setelah melewati proses otentikasi pengguna akan mendapatkan hak akses atas menu yang sesuai dengan peran dan fungsinya di dalam bank. Dikarenakan pengguna CBS berjumlah banyak, tersebar di berbagai unit kerja dengan peran dan fungsi yang beragam, maka perlu adanya metode yang sistematis dalam menentukan hak akses ini.

Organisasi bank menggunakan berbagai sistem informasi dan aplikasi komputasi untuk membantu operasional perbankan dengan banyak pengguna dari unit berbeda memiliki variasi hak akses sistem disebut *role-based access control* (**Schaad et al.**, 2001). RBAC memetakan hak akses pengguna ke dalam sebuah *User Access Matrix* (UAM) sesuai dengan perannya dalam perusahaan. Untuk menerapkan RBAC, pengguna dikelompokkan sesuai dengan perannya, lalu ditetapkan hak akses yang sesuai dengan kelompok tersebut. Pengelolaan hak akses pengguna, terkait dengan penambahan, penghapusan pengguna dan/atau perubahan hak akses akan menggunakan UAM sebagai rujukan penetapan hak akses. Hak akses dipengaruhi oleh posisi pengguna dalam struktur organisasi serta fungsi kerja rutin (**Gavrila & Barkley**, 1998). Sebagai contoh bagian transaksi memiliki hak pendebetan dengan kode transaksi, pengkreditan dengan kode transaksi, penarikan

kliring, atau tolakan masuk. Sedangkan bagian keamanan sistem memiliki hak pemeliharaan user-workstation serta pemeliharaan group menu dan program. Bank di Indonesia dapat terbagi menjadi regular dan syariah yang memiliki kantor cabang dan divisi di dalamnya. Sehingga ada banyak kemungkinan *access control* yang dimiliki pengguna.

Permasalahan muncul ketika jumlah pengguna yang dikelola cukup banyak dan variasi kebutuhan hak akses dari kelompok pengguna juga sangat beragam. Dalam kasus Bank XYZ yang menjadi fokus penelitian tesis ini, perkembangan kebutuhan menyebabkan bertambahnya variasi kebutuhan hak akses antara kelompok pengguna. Untuk itu diperlukan evaluasi pengelompokan setiap variasi hak akses kelompok pengguna atau analisis keperluan kelompok-kelompok besar pengguna dengan variasi hak akses ditetapkan sebagai kebutuhan yang berifat opsional. Penyalahgunaan hak akses memungkinkan terjadinya ancaman sehingga dalam konteks identifikasi peran pengguna juga perlu dilakukan perbandingan keseharian pemakaiannya (**Park & Giordano**, 2006). Secara konvensional ancaman umumnya baru diketahui setelah dilakukan analisis log akses pengguna pada proses audit.

Penentuan kelompok hak akses pengguna dapat dilakukan berdasarkan *top down* sebagai hasil analisis kebutuhan atau *bottom up* menggunakan data hak akses yang sudah ada seperti RoleMiner (**Vaidya et al.**, 2006)(**Vaidya et al.**, 2007) (**Vaidya et al.**, 2010). Penggalian data untuk mengenali hak akses pengguna dilakukan dengan pendekatan tanpa pembelajaran (*unsupervised*) seperti algoritma pengklasteran yang memungkinkan terjadinya overlap. RoleMiner mengatasi problem overlap dan menggunakan pendekatan *subset enumeration* serupa dengan algoritma klasik apriori atau *association rule* yang mencari pola hubungan antar data (**Agrawal et al.**, 1993). Contoh implementasi apriori adalah analisis histori belanja pelanggan untuk mengetahui sejumlah barang yang sering dibeli bersamaan sehingga pihak manajemen membuat kebijakan peletakan barang pada posisi berdekatan agar mudah dijangkau. Pada konteks RBAC, pendekatan apriori juga digunakan untuk penggalian konstrain dari hak akses (**Ma et al.**, 2012). Selain algoritma apriori yang dikategorikan pendekatan deterministik untuk eksplorasi

penggalian kelompok hak akses, terdapat pendekatan lain berbasis peluang (**Mitra et al.**, 2016) atau pembelajaran mesin (**Du & Chang**, 2014) seperti algoritma genetika (**Saenko & Kotenko**, 2017) termasuk kakas bantu RMiner untuk membantu analisis (**Li et al.**, 2013).

Kebutuhan pengelompokan hak akses pengguna pada kasus Bank XYZ memerlukan implementasi sesegera mungkin sehingga dapat dilakukan kajian kebijakan yang sesuai. RoleMiner (**Vaidya et al.**, 2006)(**Vaidya et al.**, 2007) menggunakan dataset anonim sehingga uji kebenaran data membutuhkan tahapan lebih lanjut. Sedangkan uji coba data simulasi dengan jumlah pengguna lebih sedikit memiliki kebutuhan waktu lebih lama jika diimplementasikan dengan algoritma bersifat optimasi seperti algoritma genetika (**Du & Chang**, 2014) (**Saenko & Kotenko**, 2017). Permasalahan pada penelitian disertasi ini adalah data berasal dari kasus nyata yang membutuhkan pemrosesan lebih dahulu untuk persiapan data. Selain itu jumlah pengguna serta kondisi eksisting peran pengguna yang ada sudah banyak. Oleh karena itu pada penelitian ini akan diselesaikan dengan algoritma berdasarkan pengklasteran yang sederhana namun disesuaikan dengan kondisi eksisting serta menggunakan kakas bantu tersedia. Uji coba kebenaran data dilakukan dengan diskusi lebih lanjut dengan pihak manajemen (*forum group discussion*). Hal tersebut disesuaikan dengan tujuan penelitian ini yaitu untuk mengidentifikasi, menganalisa dan menyelesaikan persoalan manajemen dan teknologi secara sistematis.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang maka permasalahan yang menjadi dalam penelitian ini adalah:

1. Bagaimana melakukan langkah-langkah persiapan data dari data asli CBS hingga menjadi fitur yang siap untuk dianalisis?
2. Bagaimana menggunakan algoritma berdasarkan pengklasteran untuk mengetahui kelompok pengguna sesuai berdasarkan peran pada kasus Bank XYZ?

3. Bagaimana melakukan evaluasi hasil pengklasteran dibanding dengan kebutuhan hak akses pengguna sehingga diberikan rekomendasi peran pengguna sesuai standar kebutuhan unit kerja Bank XYZ?

1.3 Tujuan dan Manfaat Penelitian

Tujuan penelitian ini adalah mengidentifikasi kelompok peran pengguna sesuai dengan kemiripan menu hak akses dengan pendekatan pengklasteran data pada sistem perbankan XYZ. Selanjutnya akan disusun rekomendasi standar hak akses pengguna berdasarkan hasil analisis atas kelompok peran pengguna tersebut. Diharapkan analisis kelompok hak akses tersebut dapat membantu proses monitor akses diluar kebiasaan untuk mencegah terjadinya kemungkinan penyalahgunaan hak (*fraud*).

BAB 2 TINJAUAN PUSTAKA

2.1 *Role Based Access Control* (RBAC)

Pentingnya sistem informasi dalam mendukung bisnis perusahaan menuntut adanya perlindungan keamanan yang memadai, Salah satu metode pengamanan sistem informasi adalah pengendalian hak akses. Pengendalian akses ini ditujukan untuk mengatur bahwa informasi hanya dapat diakses dan/atau diproses oleh mereka yang berhak melakukannya. Pada penerapannya, pengendalian akses mengatur pengguna atau kelompok pengguna mana saja yang dapat melakukan operasi pada suatu sumber daya.Diharapkan dengan pengendalian akses ini, kerahasiaan, integritas dan ketersediaan informasi dapat terjaga.

Salah satu metode pengendalian akses yang banyak digunakan oleh perusahaan dengan jumlah pengguna lebih dari 500 ada lah *Role Based Access Control* (RBAC). Pada metode ini, kebijakan hak akses dapat dipetakan sesuai dengan struktur organisasi perusahaan tersebut. Dalam hal ini hak akses diberikan pada seseorang sesuai dengan kebutuhan, yaitu sesuai dengan fungsi dan peran orang tersebut di perusahaan. Setiap pengguna, dalam hal ini orang yang memiliki akses ke sistem, dapat memiliki satu atau lebih peran, dan sebaliknya, setiap peran boleh jadi dapat diemban oleh lebih dari satu pengguna.Setiap peran diberikan ke pengguna sesuai tanggung jawab pekerjaannya, dan diberikan hak akses yang sesuai, yaitu data adan aplikasi apa saja yang boleh diakses untuk melaksanakan pekerjaan tersebuts

Salah satu kelebihan RBAC adalah kemudahan dalam menerapkan prinsip Least Privilege dan Separation of Duty. Prinisip Least Privilege berarti hak akses hanya diberikan secukupnya untuk dapat melaksanakan tugas atau pekerjaan pengguna. Karena RBAC disusun sesuai struktur organisasi, maka dapat dipetakan hak akses mana saja yang dapat diberikan sesuai peran pengguna.

Segregation of Duty (SOD) atau Separation of Duty adalah sebuah prinsip pengamanan dengan melakukan pemisahan tugas dan kewenangan pada sebuah proses bisnis, ke beberapa user. SOD ini umumnya dilakukan pada proses bisnis yang dianggap penting, sehingga perlu dilakukan pengendalian risiko lebih lanjut atas ancaman fraud internal. Dengan membagi tugas dan kewenangan ke lebih dari satu orang user, maka tidak ada fraud yang dapat dilakukan oleh satu orang saja. SOD menyebabkan tidak ada seseorang yang dapat melakukan dan menyembunyikan fraud. Fungsi-fungsi yang dilakukan separasi, meliputi otorisasi, pencatatan dan pengelolaan aset.

Dalam CBS, fungsi yang dianggap penting sehingga perlu dilakukan SOD meliputi fungsi otorisasi, dan input/transaksi. Seseorang yang memiliki fungsi otorisasi, tidak boleh memiliki fungsi input. Sebaliknya, seseorang yang memiliki fungsi iput tidak boleh juga memiliki fungsi otorisasi. Hal ini memastikan bahwa setiap input/transaksi yang dilakukan harus diperiksa oleh orang yang berbeda, yaitu sekurang-kurangnya orang yang memiliki kewenangan otorisasi. Dengan ini, fraud hanya dapat terjadi apabila terjadi kolusi antara pihak penginput, dengan pihak pengotorisasi.

2.2 Algoritma Penggalian Peran Pengguna

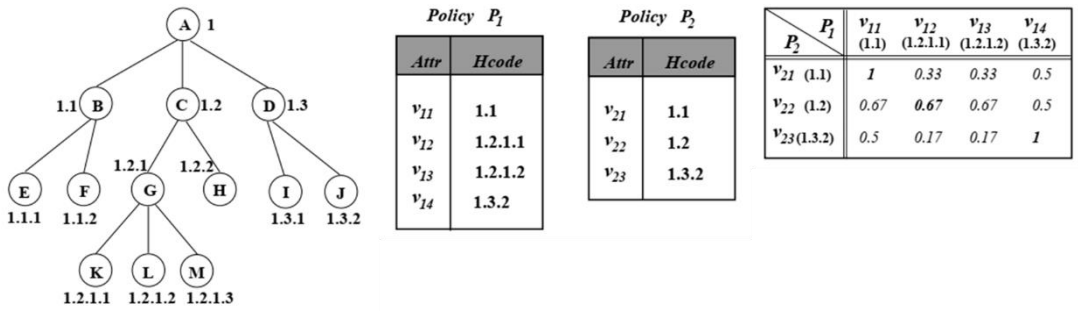
Identifikasi kelompok hak akses berdasarkan data yang telah ada (*bottom-up*) atau *role mining* umumnya dilakukan secara otomatis menggunakan teknik penggalian data (*data mining*) (**Mitra et al.**, 2016). Diawal eksplorasi identifikasi

```
for each Role i in InitRoles do
        InitRoles := InitRoles − i
        for each Role j in InitRoles do
                NewRole:= i intersection j
                if NewRole not in GenRoles then
                        Count(NewRole):= Count(NewRole) + orig_count(i)
                        Count(NewRole):= Count(NewRole) + orig_count(j)
                        Add i,j to the list of contributors for NewRole
                        GenRoles:= GenRoles union NewRole
                else
                        if i has not contributed before to NewRole then
                                Count(NewRole):= Count(NewRole) + orig_count(i)
                                Add i to the list of contributors for NewRole
                        if j has not contributed before to NewRole then
```

Gambar 2.1 Pseudocode RoleMiner

kelompok hak akses, teknik pengklasteran secara hierarkis digunakan dengan tahap persiapan data sedemikian hingga siap diproses kakas bantu (**Kuhlmann et al.**, 2003). Namun kelompok yang dihasilkan tidak memungkinan terjadinya overlap hak akses. Oleh karena itu RoleMiner (Gambar 2.1 Pseudocode RoleMinerGambar 2.1) mengatasi permasalahan duplikasi jenis hak akses dari pengguna yang memiliki peran berbeda tersebut (**Vaidya et al.**, 2006) (**Vaidya et al.**, 2010) dengan pendekatan *subset enumeration* yang mirip Algoritma Apriori. RoleMiner menggunakan dataset dengan 6000 pengguna beserta sekitar 1600 hak akses yang menghasilkan 1400 pengguna tanpa peran. Hal tersebut dapat terjadi pada kasus Bank XYZ dengan alasan yang berbeda karena hak akses dibuat saat pemelihara CBS sudah berganti serta kurang adanya dokumentasi di waktu lalu. Sehingga evaluasi kebenaran berupa diskusi dengan pihak manajemen masih dibutuhkan. Tahap persiapan data sebelum diproses kakas bantu (**Kuhlmann et al.**, 2003) juga perlu dilakukan pada kasus Bank XYZ.

Permasalahan serupa dengan hak akses adalah *policy* (**Lin et al.**, 2007) yang ditunjukkan pada Gambar 2.2. Kasus overlap pada hak akses mungkin terjadi sehingga abstraksi hierarkis dalam bentuk graf dapat membantu. Kemiripan antar peran pengguna sesuai dengan hak akses yang dimiliki akan dihitung berdasarkan pseudocode pada Gambar 2.3(**Lin et al.**, 2007). Optimasi identifikasi kelompok peran pengguna dengan abstraksi graf (**Zhang et al.**, 2007) dapat dilakukan jika pemberian hak akses mematuhi konsep hierarkis sehingga tidak terjadi overlap. Optimasi tersebut akan mengurangi jumlah kelompok peran yang perlu divalidasi pihak manajemen.



Gambar 2.2 Ilustrasi permasalahan *policy* serupa dengan hak akses

Semua pendekatan identifikasi kelompok hak akses diatas menunjukkan bahwa tidak ada formalisasi dalam penentuan kelompok peran yang benar. Bahkan pendekatan yang memperhitungkan bobot hak akses dimungkinkan karena pemberian hak umumnya mengikuti aturan hierarkis dengan sub hak cenderung memiliki bobot lebih kecil (**Ma et al.**, 2010). Algoritma yang digunakan membutuhkan proses yang sama dengan RoleMiner yaitu pendekatan Algoritma Apriori (Gambar 2.4). Aturan hierarkis juga digunakan untuk hak akses yang mempertimbangkan deskripsi unit kerja atau disebut analisis semantik di RBAC (**Molloy et al.**, 2008).

Algorithm PolicySimilarityMeasure(P_1, P_2)
Input : P_1 is a policy with n rules $\{r_{11}, r_{12}, ..., r_{1n}\}$
and P_2 is a policy with m rules $\{r_{21}, r_{22}, ..., r_{2m}\}$

1. Categorize rules in P_1 and P_2 based on their *effects*. Let $PR_1(PR_2)$ and $DR_1(DR_2)$ denote the set of permit and deny rules respectively in $P_1(P_2)$.

/* Compute similarity scores for each rule in P_1 and P_2 */
2. **foreach** rule $r_{1i} \in PR_1$
3. **foreach** rule $r_{2j} \in PR_2$
4. $S_{rule}(r_{1i}, r_{2j})$ //compute similarity score of rules
5. **foreach** rule $r_{1i} \in DR_1$
6. **foreach** rule $r_{2j} \in DR_2$
7. $S_{rule}(r_{1i}, r_{2j})$ //compute similarity score of rules

/* Compute Φ mappings */
8. $\Phi_1^P \leftarrow$ComputePhiMapping(PR_1, PR_2, ϵ)
9. $\Phi_2^P \leftarrow$ComputePhiMapping(PR_2, PR_1, ϵ)
10. $\Phi_1^D \leftarrow$ComputePhiMapping(DR_1, DR_2, ϵ)
11. $\Phi_2^D \leftarrow$ComputePhiMapping(DR_2, DR_1, ϵ)

/* Compute the rule set similarity scores */
12. **foreach** rule $r_{1i} \in P_1$
13. **if** $r_{1i} \in PR_1$ **then**
14. $rs_{1i} \leftarrow$ ComputeRuleSimilarity(r_{1i}, Φ_1^P)
15. **elsif** $r_{1i} \in DR_1$ **then**
16. $rs_{1i} \leftarrow$ ComputeRuleSimilarity(r_{1i}, Φ_1^D)
17. **foreach** rule $r_{2j} \in P_2$
18. **if** $r_{2j} \in PR_2$ **then**
19. $rs_{2j} \leftarrow$ ComputeRuleSimilarity(r_{2j}, Φ_2^P)
20. **elsif** $r_{1i} \in DR_1$ **then**
21. $rs_{2j} \leftarrow$ ComputeRuleSimilarity(r_{2j}, Φ_2^D)
22. $S_{rule-set}^P \leftarrow$ average of rs of permit rules
23. $S_{rule-set}^D \leftarrow$ average of rs of deny rules

/* Compute the overall similarity score */
24. $S_{policy}(P_1, P_2) = S_T(P_1, P_2) + w_p S_{rule-set}^P + w_d S_{rule-set}^D$

end PolicySimilarityMeasure.

Gambar 2.3 Pseudocode komputasi kemiripan *policy* serupa dengan hak akses

Pendekatan berbeda dari penentukan kelompok hak akses adalah menggunakan algoritma genetika (**Du & Chang**, 2014). Pendekatan algorita apriori adalah menggabungkan beberapa hak akses yang sering dimiliki pengguna berbeda dalam suatu kelompok yang sama. Kemudian secara iterasi dilakukan pengecekan berdasarkan kemiripan hak akses seperti yang dilakukan dengan *pseudocode* pada Gambar 2.2 dan Gambar 2.3. Sehingga pada kasus Bank XYZ yang memiliki permasalahan tersendiri antara lain belum adanya standarisasi dan dokumentasi pemberian hak akses, maka dimungkinkan terjadinya ketidakseragaman penamaan akun pengguna berdasarkan unit kerja antar cabang. Namun beberapa algoritma penggalian data diatas menunjukkan proses-proses utama antara lain penentuan

kandidat kelompok serta penentuan kemiripan antar kelompok yang memungkinkan optimasi penggabungan peran.

Require: $D \equiv (M, N, UP_{M\times N}, wminsup, F, C)$
Require: M, the number of users
Require: N, the number of permissions
Require: $UP_{M\times N}$ represents user-permission assignments
Require: $wminsup$,the weighted support threshold
Require: F represents all the frequent permission sets
Require: C represents all the candidate permission sets

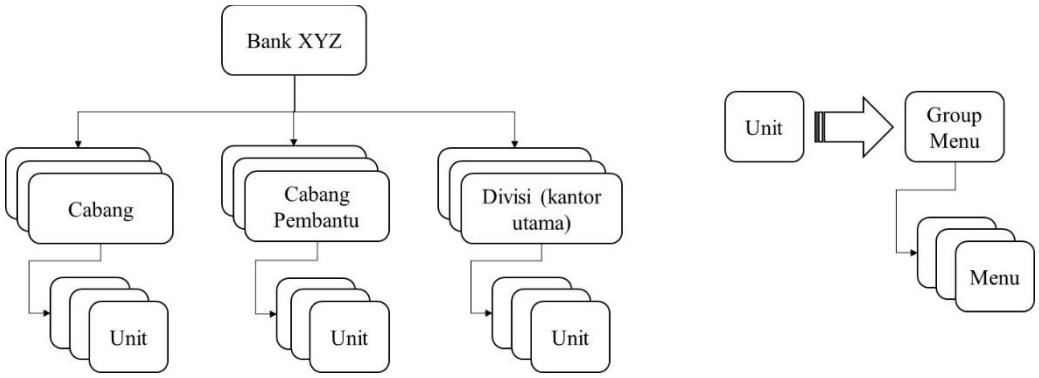
{Generate the frequent 1-permission sets}
for $(i = 1; i \leq N; i++)$ **do**
  $wsf(p_i) = w_{p_i} \times numUsers(p_i)/M$
  if $wsf(p_i) \geq wminsup$ **then**
    insert $p_i, wsf(p_i), PermUsers(p_i)$ into F_1, C_1
  else
{Generate the candidate permission sets }
    $Remove(S, w_{p_i})$
    $Sort(S)$
    for $(j = 2; j \leq maxSize; j++)$ **do**
      $maxWeight = w_{p_i} + Sum(S, j-1)$
      $minC = \lceil wminsup \times N/maxWeight \rceil$
      if $numUsers(p_i) \geq minC$ **then**
        insert $p_i, wsf(p_i), PermUsers(p_i)$ into C_1
        break
      end if
    end for
  end if
end for

{Generate the frequent k-permission sets}
for $(k = 2; C_k \neq \emptyset, k++)$ **do**
  $F_k = FrequentPermissionGen(C_{k-1}, wminsup)$
end for
$F = \cup_k F_k$
{$FrequentPermissionGen(C_{k-1}, wminsup)$}
for X and Y are in C_{k-1} **do**
  if first k-2 permissions of X and Y are the same **then**
    $PermUsers(X \cup Y) = PermUsers(X) \cap PermUsers(Y)$
    $wsf(X \cup Y) = (w_X + w_Y) \times numUsers(PermUsers(X \cup Y))/M$
    if $wsf(X \cup Y) \geq wminsup$ **then**
      insert $X \cup Y, wsf(X \cup Y), PermUsers(X \cup Y)$ into F_k, C_k
    else
      Compute $minC$ and insert the right candidate permission sets into C_k
    end if
  end if
end for

Gambar 2.4 Pseudocode identifikasi kelompok hak akses pengguna dengan bobot

2.3 Deskripsi Data

Bank XYZ adalah bank umum yang memiliki 40 cabang 200 cabang pembantu dengan 20 divisi di kantor pusat. Unit kerja utama dari bank tersebut antara lain *teller, customer service*, pelayanan nasabah, kredit, luar negeri, umum, akutansi sehingga perkiraan jumlah pengguna mencapai 4000 peran. Berdasarkan aturan tugas pokok dan fungsi telah didefinisikan hak pengguna pada sistem perbankan (Gambar 2.5). Kondisi ideal adalah suatu unit memiliki *group menu* sesuai aturan tugas pokok dengan beberapa menu dasar. Namun berdasarkan kegiatan operasional memungkinkan terjadinya penambahan hak akses seorang pengguna tanpa tercatat secara procedural dalam kebijakan tertulis. Sehingga tanpa adanya standarisasi tata kelola menyebabkan terjadinya pelanggaran hak akses.



Gambar 2.5 Ilustrasi hak akses sesuai unit kerja Bank XYZ.

Tabel 2.1 Daftar nama file hak akses pada kasus CBS Bank XYZ

No	Nama Dokumen	Keterangan
1	FILE1	Daftar nama menu yang akan menjadi hak akses seorang pengguna beserta deskripsi menu. Kolom target: nama_menu dan ket_menu 1,AKT1,MENU STAF - TRANSAKSI HARIAN,20;011;208,S06FC43T 1,AKT2,MENU STAF - TAMBAHAN,20;011;208,S06FC43T 1,AKT3,MENU STAF - INFORMASI LAYAR,20;011;208,S06FC43T 1,AKT4,MENU STAF - CETAK LAPORAN,20;011;208,S06FC43T 1,AK01,LAPORAN - BAGIAN AKUNTING,20;011;205,S06FC43T
2	FILE2	Pemetaan daftar nama pengguna beserta daftar menu yang menjadi haknya. Seorang pengguna dapat memiliki lebih dari satu menu. Kolom target: nama_user dan daftar_menu Hasil proses separasi FILE2.daftar_menu memiliki referensi ke FILE1.nama_menu TM000PW23,¢è(€€€€€€,,0,AK,CO2 GL2 GL3 HTI HTL IDEPLDEPIGRTLGRTLN3 TM000PW24,ÊÁÎ€€€€€€,,0,AK,CO2 GL2 GL3 HTI HTL IDEPLDEPIGRTLGRTLN3 TM000PW25,¢è(€€€€€€,,0,AK,CO2 GL2 GL3 HTI HTL IDEPLDEPIGRTLGRTLN3 TM000PW26,¢è(€€€€€€,,0,AK,CO2 GL2 GL3 HTI HTL IDEPLDEPIGRTLGRTLN3 TM000PW27,ACK€€€€€€€€,,0,AK,CO2 GL2 GL3 HTI HTL IDEPLDEPIGRTLGRTLN TM000PW28,BEL€€€€€€€€,,0,AK,CO2 GL2 HTI IDEPIGRTLN3 CSI TDJIKURIKY TM000PW29, åë€€€€€€€,,0,AK,CO2 GL2 GL3 HTI HTL IDEPLDEPIGRTLGRTLN3 TM000PW30,ACK€€€€€€€€,,0,AK,CO2 GL2 GL3 HTI HTL IDEPLDEPIGRTLGRTLN

Data hak pengguna pada sistem perbankan (*core banking system*, CBS) di Bank XYZ tersimpan dalam beberapa file teks berikut dengan penyebutan inisial demi menjaga kerahasiaan (Tabel 2.1). FILE1 dan FILE2 berupa dokumen teks dengan separator yaitu dokumen *csv* menyimpan data yang diilustrasikan pada Gambar 2.5. Pada penelitian ini data FILE2 diproses menjadi matriks fitur yang

kemudian dikomputasikan dengan pendekatan mesin cerdas untuk penggalian kelompok hak akses menggunakan kakas bantu.

2.4 Analisis Permasalahan Bank XYZ

Berikut adalah beberapa hal yang menjadi permasalahan pada Bank XYZ terkait hak akses pengguna.

1. Tidak ada dokumen standarisasi hak akses pengguna *Core Banking System*

Hak akses pengguna menjadi bervariasi seiring dengan perkembangan CBS karena kecenderungan pemberian akses atas setiap permintaan perubahan akun pengguna dengan alasan kebutuhan operasional di kantor cabang. Hal ini disebabkan pengelola hak akses tidak memiliki rujukan baku untuk menerima atau menolak permintaan kebutuhan tersebut. Variasi tersebut menimbulkan risiko adanya hak akses yang tidak seharusnya dimiliki oleh pengguna.

Dikarenakan tidak adanya dokumentasi maka terdapat indikasi hak akses pengguna yang tidak seharusnya. Pengikutsertaan data tersebut akan berpotensi menimbulkan pengelompokan peran yang tidak sesuai. Oleh karena itu perlu dilakukan penyaringan data untuk meningkatkan akurasi pengelompokan agar dapat mengidentifikasi pemberian hak akses yang tidak tepat.

2. Variasi peran dalam organisasi Bank XYZ terkait perumusan standarisasi hak akses pengguna

Terdapat sekitar 200 cabang dan cabang pembantu di organisasi Bank XYZ dengan berbagai fungsi sesuai kebutuhan bisnis cabang. Pengelompokan hak akses sesuai dengan peran akan menghasilkan jumlah kelompok pengguna yang terlampau besar untuk dikelola dan berpotensi sering berubah seiring dinamika bisnis. Akan tetapi jika beberapa kelompok tersebut digabungkan maka timbul potensi berkurangnya efektifitas standarisasi hak akses karena pemberian hak akses yang tidak seharusnya. Oleh karena itu perlu dilakukan pengelompokan hak akses yang optimal bagi Bank XYZ.

Berdasarkan kegiatan operasional pada kantor cabang dan cabang pembantu yang relatif kecil dimungkinkan adanya pengguna yang memiliki peran

rangkap. Hal tersebut menambahkan kompleksitas pengelompokkan pengguna. Kesulitan identifikasi diakibatkan pengguna dengan peran rangkap memiliki peran utama sesuai dengan kodifikasi penamaan, namun memiliki sejumlah menu sesuai hak akses untuk peran lain yang dirangkap. Diharapkan dengan penelitian ini dapat diidentifikasi pengguna dengan peran rangkap serta pengelompokan hak akses yang efektif.

3. Kebutuhan klasifikasi pemberian hak akses

Kategori pertama adalah hak akses yang harus dimiliki oleh sebuah kelompok peran. Hak akses yang masuk kategori pertama diberikan secara default ke pengguna sesuai dengan kelompok peran. Kategori kedua adalah hak akses opsional sesuai permintaan pengguna di suatu kelompok peran. Ketiadaan hak akses bersifat opsional tidak akan mengganggu operasional pengguna yang tidak membutuhkan hak akses tersebut. Sedangkan pengguna yang membutuhkan hak akses opsional dapat terakomodir sehingga tidak mengganggu operasional bank. Kategori ketiga adalah hak akses bersifat terbatas yang secara default tidak boleh dimiliki oleh suatu kelompok peran. Pemberian hak akses kategori terbatas harus melalui prosedur yang sangat ketat dan dipantau setiap penggunaannya. Pengelola hak akses secara berkala akan melakukan review atas hak akses kategori terbatas untuk meminimalkan risiko pelanggaran hak akses.

2.5 Pengklasteran Kelompok Hak Akses

Pengklasteran adalah membagi populasi ke dalam sejumlah kelompok sedemikian hingga titik data dalam kelompok yang sama memiliki lebih banyak kemiripan daripada titik data pada kelompok lain. Pengklasteran kelompok hak akses pada penelitian ini merupakan sebuah pendekatan *bottom-up* untuk mengidentifikasi kelompok peran pengguna sesuai dengan data hak akses yang dimiliki (**Vaidya et al.**, 2010). Pendekatan ini ditempuh karena ketiadaan informasi rujukan berupa dokumen standarisasi hak akses pengguna.

Dari beberapa algoritma pengklasteran yang ada, salah satu yang dapat digunakan untuk mengidentifikasi kelompok hak akses adalah algoritma dengan menggunakan model centroid. Pada model ini, hak akses dibagi kedalam sejumlah

kelompok/klaster berdasarkan kedekatan titik data yang merepresentasikan hak akses pengguna, dengan titik tengah (centroid) kluster tersebut. Pengklasteran dilakukan secara iteratif hingga diperoleh jarak rata-rata terdekat dari semua titik data pada klaster dengan centroidnya.

Penentuan keanggotaan sebuah titik data yang merepresentasikan hak akses pengguna, memerlukan perbandingan antara kedekatan jarak titik data tersebut dengan centroid, terhadap jarak antar centroid. Apabila jarak dari titik data ke centroidnya jauh lebih dekat dibandingkan jarak antar centroid (merepresentasikan jarak antar klaster), maka identifikasi keanggotaan klaster relatif lebih mudah dilakukan karena sebaran titik data akan cenderung mengumpul dekat centroid (*scatter within*). Sebaran titik data yang tidak mengumpul dekat centroid selanjutnya dapat diidentifikasi sebagai klaster baru. Sebaliknya, apabila jarak dari titik data ke centroidnya lebih besar dibandingkan jarak antar centroid, maka identifikasi keanggotaan klaster akan lebih sulit dilakukan karena kemungkinan adanya sebaran titik data yang dekat dengan lebih dari satu centroid (*scatter between*).

Pada penelitian ini, centroid merepresentasikan menu hak akses yang seharusnya dimiliki oleh pengguna dengan peran tertentu. Dengan kata lain, jumlah kluster akan menunjukkan jumlah peran pengguna yang ada dalam CBS. Hasil pengelompokan hak akses ini selanjutnya perlu divalidasi lebih lanjut. Validasi klaster ini berupa pengukuran seberapa baik hasil pengklasteran. Validasi pertama adalah apakah kecenderungan hasil pengklasteran cocok dengan permasalahan yang dihadapi, yaitu apakah anggota kelompok hak akses yang ada dalam klaster cukup memiliki kemiripan satu sama lain. Validasi kedua terkait jumlah klaster, apakah telah mencerminkan jumlah variasi peran pengguna. Validasi ketiga meliputi berbagai metode evaluasi untuk mengukur hasil klaster, seperti *Dunn index* dan *Silhouette coefficient*.

2.6 Algoritma Pengklasteran K-Means

Pseudocode pengklasteran K-Means yang akan menjadi dasar algoritma pada penelitian ini ditunjukkan pada Gambar 2.6 (**Knox**, 2018). Setiap data atau

datum yang akan dikelompokkan adalah seorang nama_user dari FILE2 (Tabel 2.1). Sebuah data dianggap sebagai suatu vektor kolom x_i dari koleksi pengguna $i = 1 \ldots n$ yang memiliki dimensi m sesuai dengan jumlah nama_menu pada FILE1. Terdapat sekitar 450 menu pada FILE1 sehingga $m = 450$ namun jumlah dimensi tersebut dapat berkurang menurut konteks masalah yang akan diuraikan pada bagian metode penelitian.

Permasalahan pertama pada algoritma K-Means adalah penentuan representasi kelompok atau *cluster center* atau *centroid* yang umumnya ditentukan secara random. Namun pada kasus Bank XYZ akan dilakukan analisis data awal sehingga direncanakan beberapa pengguna dengan peran tertentu sesuai unit kerja. Penjelasan tersebut akan diuraikan pada bagian metode penelitian.

Permasalahan kedua pada algoritma K-Means adalah penentuan kemiripan antar pengguna. Umumnya suatu pasangan data x_i dan x_j bisa dimasukkan dalam kelompok yang sama apabila memiliki jarak terdekat misal berdasarkan jarak Euclidean $eucl(x_i, x_j) = \sqrt{\sum_{k=1\ldots m}(x_{ik} - x_{jk})^2}$. Namun pada kasus Bank XYZ kemiripan hak akses pengguna tidak bisa dihitung dengan jarak Euclidean.

(1) Given a set of putative cluster centers $\widehat{\mu}_1, \ldots, \widehat{\mu}_k$, assign each datum to the cluster which has the nearest center: for $i = 1, \ldots, n$, datum x_i is assigned to cluster $\operatorname{argmin}_{j=1,\ldots,k} \|x_i - \widehat{\mu}_j\|$.

(2) Set the center of each cluster to be the mean of the data in the cluster: for $j = 1, \ldots, k$,

$$\widehat{\mu}_j = \frac{1}{\text{number of data in cluster } j} \sum_{x_i \in \text{ cluster } j} x_i \ .$$

Gambar 2.6 Pseudocode Pengklasteran K-Means

Sebagai contoh hak akses x_i adalah ['LN1', 'LN2', 'LN3', 'LN4', 'GJ01'] sedangkan x_j memiliki hak akses ['LN3', 'LN4', 'LN9', 'GL2', 'CSI']. Maka untuk menghitung kemiripan pengguna x_i dan x_j dengan indeks Jaccard (1) (**Knox**, 2018).

Pada kasus tersebut, $x_i \cap x_j = 2$ untuk ['LN3', 'LN4'] dan $x_i \cup x_j = 8$ untuk ['LN1', 'LN2', 'LN3', 'LN4', 'GJ01'] + ['LN3', 'LN4', 'LN9', 'GL2', 'CSI'] - ['LN3', 'LN4'].

$$sim(x_i, x_j) = \frac{x_i \cap x_j}{x_i \cup x_j} \quad (1)$$

Namun jika unsur konteks diperhatikan maka menu ['LN3', 'LN4', 'LN9'] dapat dianggap sama sebagai menu ['LN'] sehingga nilai indeks Jaccard dapat berbeda. Hal tersebut dapat dilakukan berdasarkan konsep semantik (**Molloy et al.**, 2008) serta bobot (**Ma et al.**, 2010) pada RBAC. Penjelasan tersebut akan diuraikan pada bagian metode penelitian.

2.7 Algoritma Pengklasteran Fuzzy C-Means

Selain algoritma K-Means juga dilakukan uji coba dengan Fuzzy C-Means (FCM) (**Dunn**, 1973) (**Bezdek**, 1981) untuk mengatasi masalah overlap data pengguna karena setiap user x_i menjadi anggota klaster j dengan nilai keanggotaan tertentu $\mu_j(x_i)$. Perbedaan K-Means dengan FCM adalah *soft-clustering vs hard-clustering* artinya suatu data dimungkinkan menjadi anggota lebih dari satu klaster berdasarkan nilai probabilitas tertentu.

Langkah-langkah dalam Algoritma FCM adalah sebagai berikut:

1. Inisialisasi matriks keanggotaan U dengan ukuran baris sejumlah data dan kolom sejumlah klaster $= 1 \ldots C$. Total nilai keanggotaan suatu data $u_{ij} = \mu_j(x_i)$ adalah 1.

$$\sum_{j=1}^{C} \mu_j(x_i) = 1$$

2. Hitung centroid setiap kluster dengan nilai parameter fuzifikasi $1.25 \leq m \leq 2$ dengan N_j adalah jumlah data yang masuk dalam klaster j. Jika $m = 1$ maka data tidak akan overlap dan hanya masuk dalam satu klaster. Jika $m > 1$ maka data dapat overlap dan masuk ke banyak klaster. Jika suatu data memiliki dimensi k maka x_i adalah vektor berukuran $1 \times k$ sehingga satu elemen dalam vektor dinotasikan sebagai x_{ik}. Maka centroid C_j juga berupa vektor berukuran $1 \times k$ dan setiap elemen nya dinotasikan sebagai C_{jk}.

$$C_{jk} = \frac{\sum_{i=1}^{N} u_{ij}^{m} \times x_{ik}}{\sum_{i=1}^{N} u_{ij}^{m}}$$

3. Hitung disimilaritas antar data x_i dengan centroid C_j menggunakan jarak Euclidean d_{ij}.
4. Perbarui matriks keanggotaan U dengan setiap nilai elemennya.

$$u_{ij} = \mu_j(x_i) = \frac{\left(1/d_{ij}\right)^{1/m-1}}{\sum_{l=1}^{C}\left(1/d_{il}\right)^{1/m-1}}$$

5. Ulang langkah 2-4 sampai konvergen yaitu selisih nilai matriks keanggotaan pada iterasi ke-n dibanding dengan iterasi sebelumnya n-1 tidak melebihi suatu nilai ambang.

$$U^n - U^{n-1} < threshold$$

Pada permasalahan rekomendasi kebijakan standarisasi menu hak akses dimungkinkan satu pengguna memiliki peran berbeda selain peran utamanya. Oleh karena itu dengan penerapan FCM pada data-data menu dari pengguna dapat dilihat kemungkinan pemetaan multi peran.

2.8 Algoritma Pengklasteran DBSCAN

Algoritma pengklasteran DBSCAN (*Density-based spatial clustering of applications with noise*) mengelompokkan data yang memiliki jarak berdekatan berdasarkan dan memenuhi minimal jumlah data atau disebut juga memiliki kepadatan (*density* atau densitas) lebih tinggi (**Ester et al.**, 1996). Sehingga data-data yang jarang dengan kepadatan rendah disebut sebagai data *outlier*. Parameter DBSCAN adalah *eps* sebagai nilai ambang antar data agar dikatakan berdekatan dan *minPoints* sebagai jumlah minimal data dalam suatu klaster.

Algoritma pengklasteran DBSCAN (Gambar 2.7) dilakukan secara iterasi ke semua data pengguna. Untuk setiap data yang berdekatan dengan minimal jarak *eps* maka data-data tersebut akan menjadi satu kelompok. Jika syarat tersebut tidak terpenuhi maka data-data dijadikan sebagai data outlier.

Pada penelitian ini digunakan Python Library Scikit-learn (**Pedregosa et al.**, 2011) untuk algoritma pengklasteran K-Means, FCM, dan DBSCAN untuk mengelompokkan pengguna berdasarkan menu hak akses yang dimiliki. Jika terdapat pengguna di daerah yang padat dalam suatu klaster dan hak akses menu dari pengguna menunjukkan peran berbeda, maka perlu dilakukan rekomendasi standarisasi hak akses atas peran tersebut.

```
DBSCAN (SetOfPoints, Eps, MinPts)
// SetOfPoints is UNCLASSIFIED
  ClusterId := nextId(NOISE);
  FOR i FROM 1 TO SetOfPoints.size DO
    Point := SetOfPoints.get(i);
    IF Point.ClId = UNCLASSIFIED THEN
      IF ExpandCluster(SetOfPoints, Point,
            ClusterId, Eps, MinPts) THEN
        ClusterId := nextId(ClusterId)
      END IF
    END IF
  END FOR
END; // DBSCAN
```

```
ExpandCluster(SetOfPoints, Point, ClId, Eps,
              MinPts) : Boolean;
  seeds:=SetOfPoints.regionQuery(Point,Eps);
  IF seeds.size<MinPts THEN // no core point
    SetOfPoint.changeClId(Point,NOISE);
    RETURN False;
  ELSE   // all points in seeds are density-
         // reachable from Point
    SetOfPoints.changeClIds(seeds,ClId);
    seeds.delete(Point);
    WHILE seeds <> Empty DO
      currentP := seeds.first();
      result := SetOfPoints.regionQuery(currentP,
                                        Eps);
      IF result.size >= MinPts THEN
        FOR i FROM 1 TO result.size DO
          resultP := result.get(i);
          IF resultP.ClId
              IN {UNCLASSIFIED, NOISE} THEN
            IF resultP.ClId = UNCLASSIFIED THEN
              seeds.append(resultP);
            END IF;
            SetOfPoints.changeClId(resultP,ClId);
          END IF; // UNCLASSIFIED or NOISE
        END FOR;
      END IF; // result.size >= MinPts
      seeds.delete(currentP);
    END WHILE; // seeds <> Empty
    RETURN True;
  END IF
END; // ExpandCluster
```

Gambar 2.7 Pseudocode Pengklasteran DBSCAN

2.9 Algoritma t-SNE untuk Transformasi Data

t-Distributed Stochastic Neighbor Embedding (t-SNE) dapat digunakan untuk reduksi data multi dimensi menjadi dua dimensi untuk kemudahan visualisasi data (**van der Maaten & Hinton**, 2008). Reduksi dimensi dengan t-SNE memiliki kelebihan yang tetap mempertahankan struktur data aslinya (Gambar 2.8). t-SNE meghitung kemiripan antar data berdimensi tinggi yang kemudian ditransformasi menggunakan Gaussian distribution dan Student t-distribution.

Data: data set $\mathcal{X} = \{x_1, x_2, ..., x_n\}$,
cost function parameters: perplexity *Perp*,
optimization parameters: number of iterations T, learning rate η, momentum $\alpha(t)$.
Result: low-dimensional data representation $\mathcal{Y}^{(T)} = \{y_1, y_2, ..., y_n\}$.
begin
  compute pairwise affinities $p_{j|i}$ with perplexity *Perp* (using Equation 1)
  set $p_{ij} = \frac{p_{j|i} + p_{i|j}}{2n}$
  sample initial solution $\mathcal{Y}^{(0)} = \{y_1, y_2, ..., y_n\}$ from $\mathcal{N}(0, 10^{-4}I)$
  for *t=1* **to** T **do**
    compute low-dimensional affinities q_{ij} (using Equation 4)
    compute gradient $\frac{\delta C}{\delta \mathcal{Y}}$ (using Equation 5)
    set $\mathcal{Y}^{(t)} = \mathcal{Y}^{(t-1)} + \eta \frac{\delta C}{\delta \mathcal{Y}} + \alpha(t)\left(\mathcal{Y}^{(t-1)} - \mathcal{Y}^{(t-2)}\right)$
  end
end

Gambar 2.8 Pseudocode transformasi t-SNE

Pada penelitian ini transformasi t-SNE juga menggunakan Python Library Scikit-learn (**Pedregosa et al.**, 2011).

BAB 3 METODE PENELITIAN

3.1 Persiapan Data

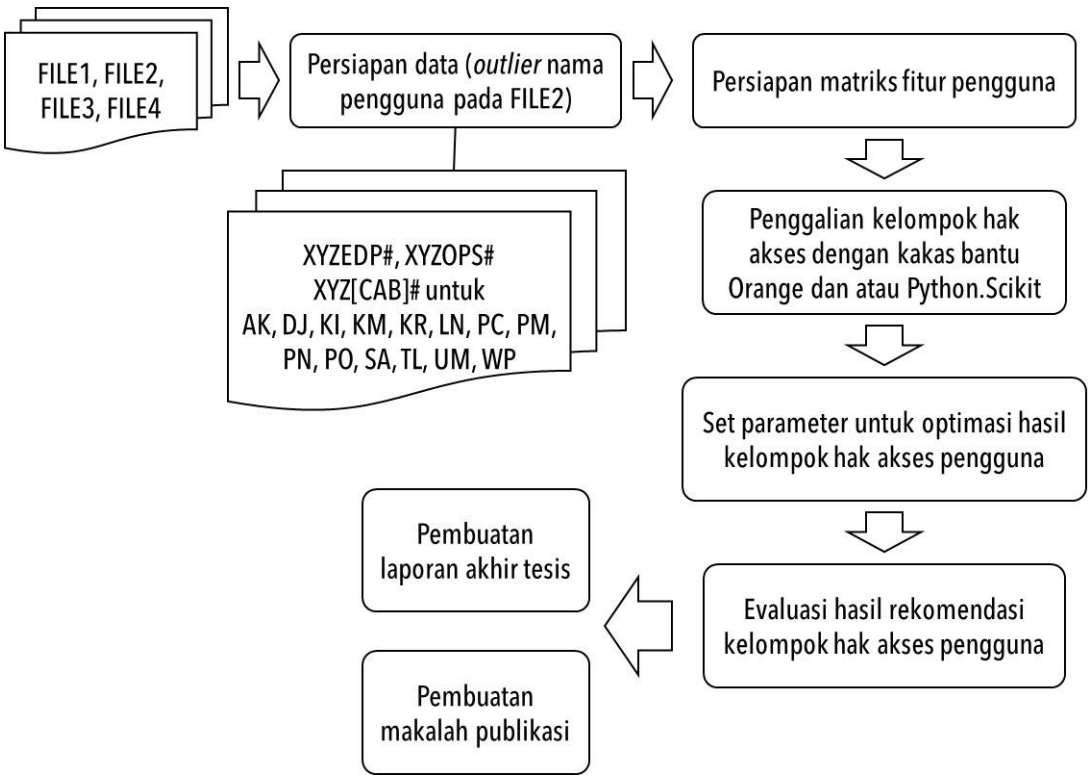
Terdapat sekitar 7500 data pengguna beserta hak akses menu pada FILE2 (Tabel 2.1) yang akan diproses pada penelitian ini. Namun fokus masalah akan diutamakan pada pengguna dengan pola penaaman FILE2.nama_user seperti yang ditunjukkan pada Tabel 3.1.2 Permasalahan dicurigai adanya duplikasi hak akses menu untuk pengguna XYZEDP# dengan XYZOPS# sehingga penelitian ini dilakukan.

Tabel 3.1 Pola nama pengguna pada kasus Bank XYZ

No	Pola nama_user	Keterangan	Jumlah Pengguna
1	XYZEDP#	Pengguna adalah staf cabang utama yang bertanggung jawab pada bagian teknologi informasi	268
2	XYZOPS#	Pengguna adalah staf cabang utama yang bertanggung jawab pada bagian teknologi informasi khususnya operasional seperti pelayanan ATM dan lainnya	52
3	XYZ[CAB]# Variasi kode cabang: 001-199	Pengguna adalah staf cabang utama maupun daerah sesuai dengan peran seperti XYZ001AK23 adalah pengguna pada cabang utama dengan kode 001 untuk peran Akutansi (AK)	6365

Sedangkan peran pengguna di kantor cabang dengan pola XYZ[CAB]# memiliki variasi seperti yang ditunjukkan pada Tabel 3.1. Tidak semua cabang memiliki peran tersebut. Berdasarkan jumlah pengguna di cabang Bank XYZ terdapat beberapa nama yang akan diabaikan datanya seperti XYZ[CAB]PC karena hanya ada seorang pimpinan cabang sehingga duplikasi hak akses tidak dimungkinkan. Analisis lain yaitu mengabaikan akun XYZ026DM01 yang hanya dimiliki oleh seorang pengguna. Hal itu mungkin terjadi karena tidak adanya standarisasi dokumentasi hak akses sehingga penelusuran latar belakang pembuatan

akun untuk semua cabang. Selain itu terdapat sekitar 950 akun pengguna yang tidak memiliki hak akses menu dengan sekitar 670 akun diantaranya termasuk peran TO dan TL. Hal tersebut terjadi karena kurangnya dokumentasi dan standarisasi proses pembuatan akun sehingga permasalahan tersebut menjadi latar belakang penelitian ini untuk membantu rekomendasi kebijakan dan standarisasi pengaturan hak akses pengguna CBS.



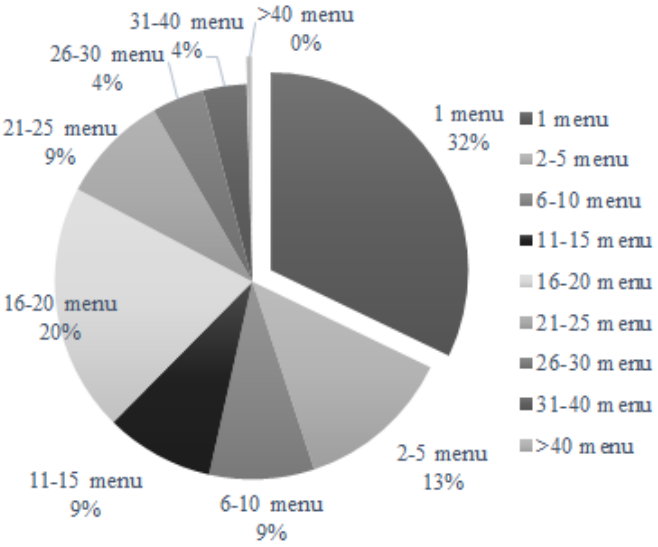
Gambar 3.1 Diagram alir pengklasteran kelompok hak akses pengguna berdasarkan peran

Hal-hal berikut menjadi perhatian untuk pengklasteran kelompok hak akses pengguna:

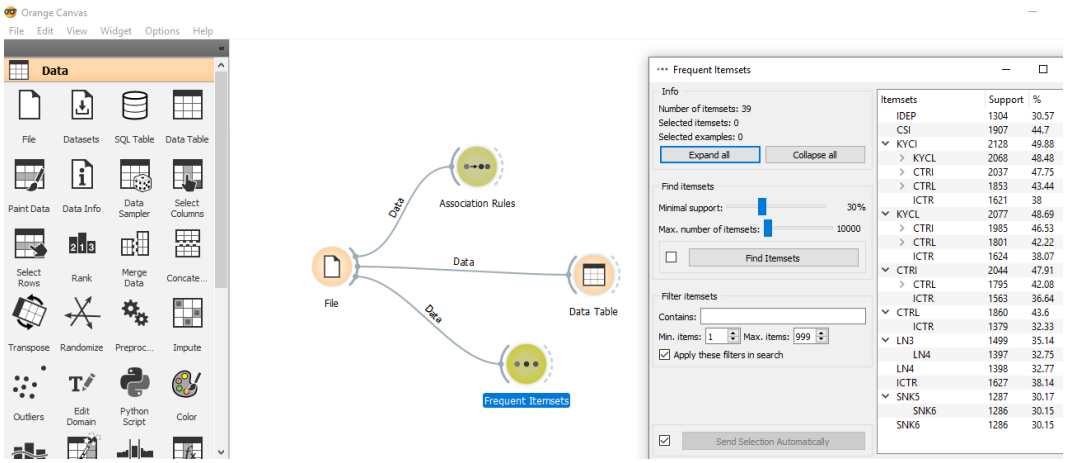
1. Reduksi dimensi data

Data seorang akun pengguna x_i berupa vektor dengan jumlah dimensi adalah 469 menu yang tersedia pada FILE1. Tidak semua menu digunakan karena hanya beberapa peran pengguna yang akan diproses seperti yang ditunjukkan pada Gambar 3.1. Analisis menu yang sering dimiliki oleh setiap peran setidaknya 10% dari jumlah pengguna ditunjukkan pada Tabel 3.3. Sebagai contoh untuk peran

Analisis data pengguna berdasarkan jumlah menu terlihat pada Gambar 3.3. Sebagian besar pengguna memiliki menu terbatas yaitu < 10 menu. Hanya sebagian kecil pengguna dengan kemungkinan peran ganda yang diindikasikan dari banyaknya menu.



Gambar 3.3 Statistik pengguna dengan jumlah menu yang dimiliki



Gambar 3.4 Contoh penggunaan kakas bantu Orange untuk algoritma Apriori

tampilan dalam tabel, tampilan set menu yang sering muncul dan tampilan aturan yang disarankan. Untuk analisis lebih lanjut, set menu yang sering muncul pada Gambar 3.4 (CSI, CTRI, CTRL, ICTR, IDEP, KYCI, KYCL, LN3, LN4, SNK5, SNK6) ternyata dimiliki oleh hampir semua peran pengguna pada Gambar 3.5. Sedangkan potongan aturan yang direkomendasikan oleh kakas bantu Orange ditampilkan pada Gambar 3.6. Namun banyaknya jumlah aturan yang direkomendasikan serta adanya duplikasi aturan sehingga dimungkinkan untuk penggabungan membuat Algoritma Apriori membutuhkan pemrosesan lebih lanjut. Analisis penggalian kelompok hak akses pengguna selanjutnya akan dilakukan dengan pendekatan pengklasteran.

2. Penentuan *centroid* awal kelompok hak akses

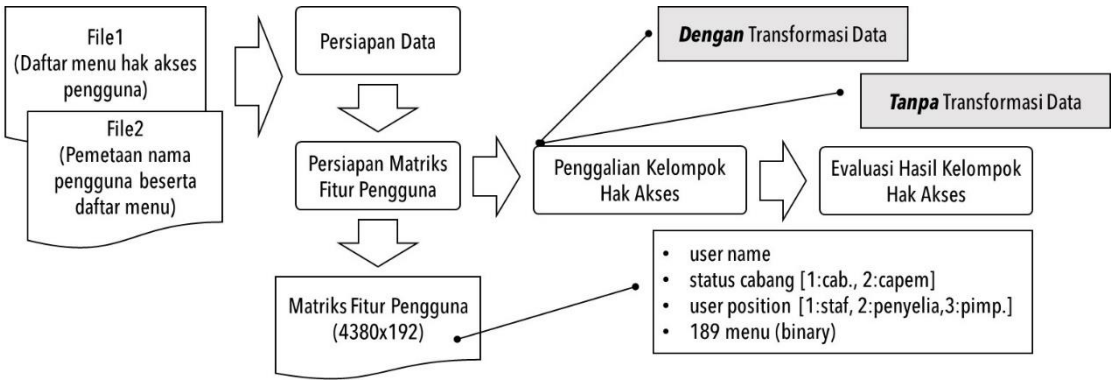
Penentuan *centroid* dapat menggunakan informasi menu utama yang ditunjukkan pada Tabel 3.3. Diasumsikan bahwa *centroid* adalah data akun pengguna yang memiliki sebagian besar hak akses menu. Jumlah kelompok ditetapkan sebanyak 14 kelompok mengikuti konteks peran pengguna seperti uraian berikut.

3. Memperhitungkan konteks peran dari hak akses

Standarisasi penamaan akun yang disesuaikan dengan peran pengguna sudah dilakukan pada kasus CBS Bank XYZ seperti yang ditunjukkan pada Tabel 3.3. Namun pada pelaksanaannya terdapat kemungkinan pemberian hak akses menu yang melebihi peran pengguna seharusnya. Oleh karena itu penelitian ini dilakukan. Jika konteks peran tersebut dianggap sebagai kelas maka data input dapat diproses dengan pendekatan penggalian data dengan pembelajaran (*supervised*). Terdapat 14 kelas yang akan ditetapkan yaitu pada Tabel 3.3 serta dua kelas EDP dan OPS pada Tabel 2.1. Dikarenakan terdapat ketidakseimbangan dalam jumlah data maka pendekatan *K-fold validation* sebagai salah satu cara proses random untuk pengambilan sampel data akan dilakukan.

3.2 Pengklasteran Menu Pengguna

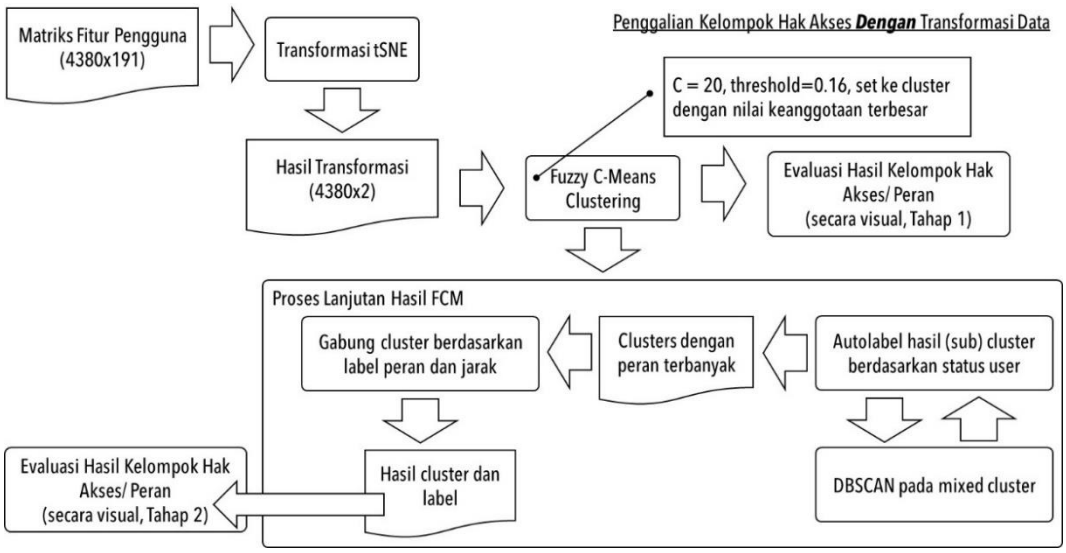
Pada penelitian ini dilakukan dua pendekatan pengklasteran yaitu dengan dan tanpa transformasi data menu pengguna menjadi dua dimensi untuk kemudahan dalam analisis hasil (Gambar 3.7).



Gambar 3.7 Pengklasteran menu untuk rekomendasi kelompok hak akses

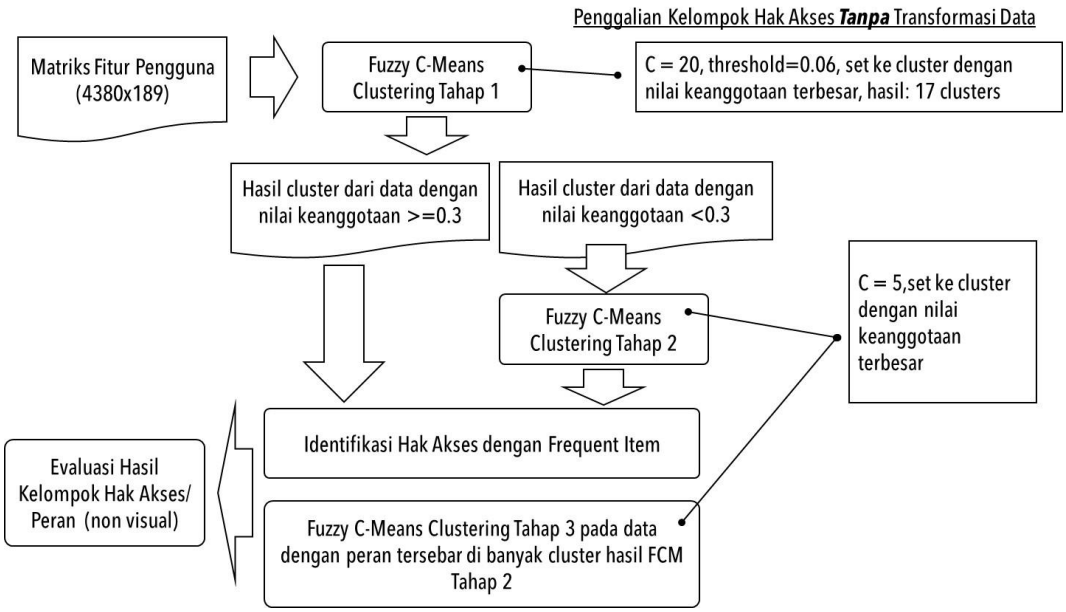
Matriks fitur pengguna dari hasil proses FILE1 dan FILE2 berukuran 4380x192 untuk 4380 data pengguna dengan fitur sebagai berikut: nama pengguna, status pengguna (1:staf, 2:penyelia,3:pimpinan), status kantor pengguna (1:cabang, 2:cabang pembantu), serta 189 menu hasil reduksi. Kemudian pengklasteran pada matriks fitur dilakukan melalui dua pendekatan:

a. Pengklasteran hak akses dengan transformasi t-SNE (Pendekatan 1). Pendekatan ini dilakukan untuk mengetahui kelompok hak akses dengan menggunakan transformasi t-SNE (Gambar 3.8) sebelum pengklasteran untuk kemudahan validasi secara visual. Pendekatan 1 menggunakan beberapa tahap pengklasteran yaitu FCM dan DBSCAN. Proses pengklasteran kedua dengan DBSCAN hanya dilakukan pada klaster tertentu yang disebut dengan klaster campuran. Kemudian hasil klaster akan dicek lebih lanjut untuk menentukan penggabungan klaster sesuai syarat tertentu.



Gambar 3.8 Pengklasteran kelompok hak akses dengan transformasi data t-SNE

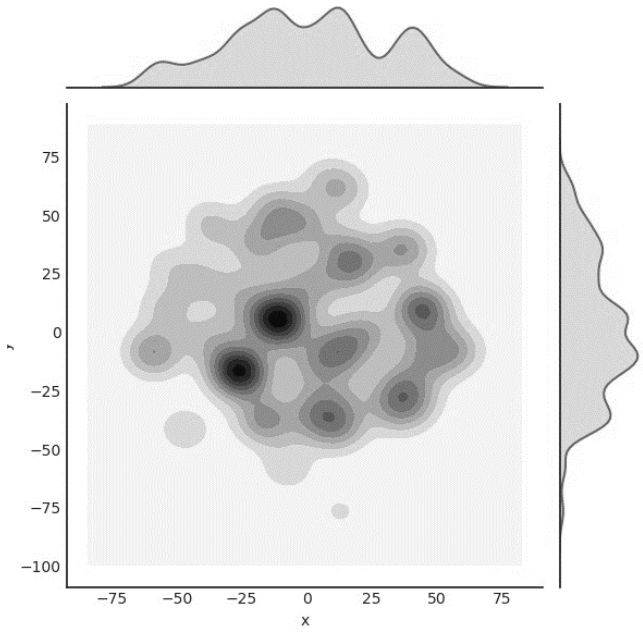
b. Pengklasteran hak akses dengan transformasi t-SNE (Pendekatan 2). Pendekatan ini dilakukan dengan melakukan pengklasteran tanpa melalui transformasi t-SNE (Gambar 3.9). Pendekatan 2 memiliki beberapa tahap pengklasteran dari FCM.



Gambar 3.9 Pengklasteran kelompok hak akses tanpa transformasi data t-SNE

3.2.1 Hasil Pengklasteran dengan Pendekatan 1

Setelah dilakukan transformasi t-SNE pada matriks fitur pengguna dengan Python Library sklearn.manifold (**Pedregosa et al.**, 2011) maka setiap pengguna dapat dipetakan pada tampilan dua dimensi Gambar 3.10 Visualisasi data pengguna setelah transformasi t-SNEGambar 3.10. Reduksi dimensi dilakukan dari 191 (189 menu, status pengguna dan status kantor pengguna) kolom menjadi dua kolom (nilai x dan y untuk visualisasi). Terlihat bahwa setidaknya terdapat >10 peran utama berdasarkan kepadatan data. Kemudian FCM dilakukan pada matriks dua dimensi dengan variasi jumlah klaster C ditunjukkan di bahasan hasil. Sebagai pembanding, pada matriks pengguna dua dimensi dilakukan pengklasteran dengan algoritma DBScan, KMeans++, FCM, dan FCM dengan modifikasi DBScan. Pada proses selanjutnya digunakan data dua dimensi dan peran pengguna yang diambil dari nama pengguna.



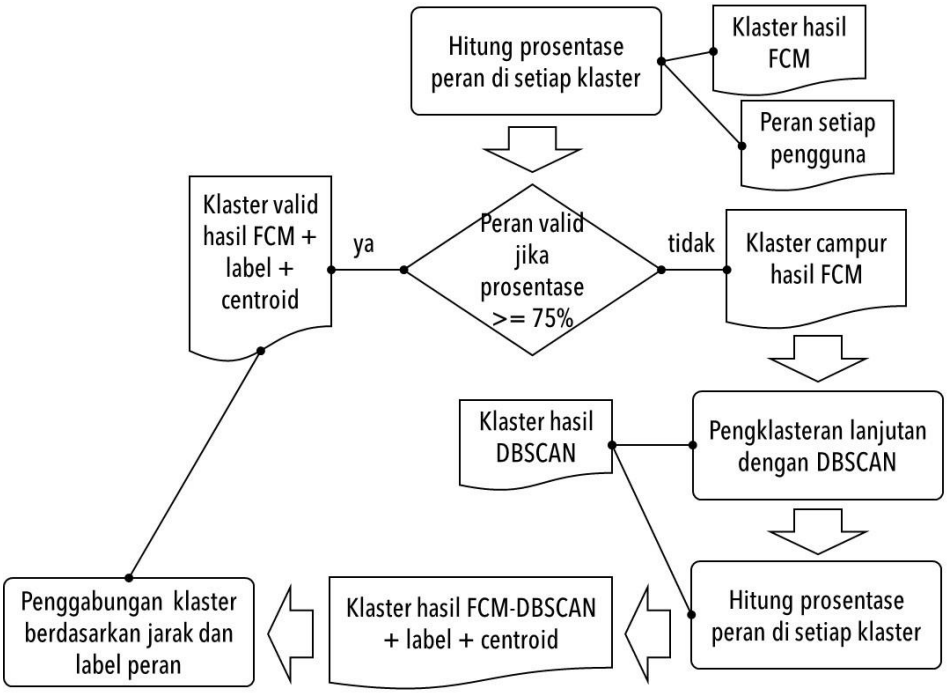
Gambar 3.10 Visualisasi data pengguna setelah transformasi t-SNE

Kemudian proses selanjutnya adalah penentuan pengklasteran lanjutan dengan DBSCAN serta pemberian auto-label berdasarkan peran terbanyak di suatu

Total Pengguna 166

Setelah proses DBSCAN dari hasil FCM maka setiap sub klaster akan memiliki auto-label dengan prosentase peran >60%. Kondisi tersebut dikatakan sub klaster memiliki peran yang valid. Jika kondisi tersebut tidak terpenuhi maka label ditentukan dari peran sub klaster terdekat berdasarkan jarak antar centroid. Semua proses setelah FCM dari transformasi data t-SNE ditunjukkan pada Gambar 3.11.

Sebagai contoh hasil FCM dari matriks pengguna dengan C=20 memiliki delapan klaster dengan peran valid dan 12 klaster campur yang memerlukan pengklasteran lanjutan DBSCAN. Deskripsi hasil dan pembahasan lebih lengkap untuk pengklasteran kelompok hak akses tanpa transformasi data t-SNE yang disebut Pendekatan 1 diuraikan pada Bab 4.



Gambar 3.11 Modifikasi FCM untuk pengklasteran kelompok hak akses dengan transformasi data t-SNE

3.2.2 Hasil Pengklasteran dengan Pendekatan 2

Pengklasteran tanpa transformasi data t-SNE dan langsung menggunakan 189 menu dilakukan setelah analisis hasil akhir pengklasteran diatas menunjukkan kelompok hak akses yang masih tercampur. Deskripsi hasil dan pembahasan dari Pendekatan 1 diuraikan pada Bab 4. Kemudian pada sub bahasan ini dilakukan Pendekatan 2 yaitu pengklasteran tanpa transformasi data t-SNE. Algoritma FCM tetap dilaksanakan pada proses awal dengan data 189 dimensi dengan nilai C=20. Sebagai catatan kodifikasi klaster C1-C20 pada Pendekatan 2 tidak menunjukkan hasil yang sama dengan hasil dari Pendekatan 1. Data seorang pengguna pada kedua pendekatan tersebut hanya akan dipetakan ke satu klaster tertentu berdasarkan nilai keanggotaan terbesar. Pada analisis awal hasil klaster FCM berdasarkan menu terlihat jika suatu data memiliki nilai keanggotaan terbesar >= nilai ambang 0.3 maka peran yang dimiliki sudah sesuai dengan hak akses menu. Informasi peran diambil dari nama pengguna sedemikian hingga xxx001KM01 akan memiliki peran KM yang dianggap memiliki tugas pemasaran kredit di bank.

Meskipun FCM berdasarkan menu memiliki parameter awal C=20, namun ternyata hanya ada 13 klaster yang terbentuk berdasarkan nilai keanggotaan terbesar (uraian lebih lengkap ada pada Bab 4). Pada Pendekatan 1 klaster valid setelah FCM jika memiliki peran dengan prosentase terbesar. Namun pada Pendekatan 2 informasi peran tidak digunakan sehingga klaster valid adalah klaster hasil FCM dengan nilai keanggotaan terbesar >= 0.3. Dari 13 klaster hasil FCM terdapat 7 klaster yang memiliki klaster valid dan tidak valid sehingga perlu diklaster ulang sebagian datanya. Selain itu ada 2 klaster yang harus diklaster ulang semuanya karena tidak memiliki data dengan nilai keanggotaan terbesar >= 0.3 serta 4 klaster yang akan dianggap outlier karena tidak valid dan hanya memiliki <10 anggota. Proses pengklasteran lanjutan dilakukan menggunakan FCM pada data-data di setiap klaster jika nilai keanggotaan terbesar < 0.3 dengan C=5. Kemudian dilakukan analisis menu pada hasil sub klaster untuk rekomendasi kebijakan standarisasi menu hak akses.

3.3 Evaluasi Hasil Pengklasteran Menu Pengguna

Untuk menilai hasil pengelompokan peran dilakukan pengukuran atas tingkat kesalahan pelabelan dan tingkat kemiripan menu hak akses pengguna dalam klaster yang dihasilkan.

Tabel 3.5 Contoh Perhitungan Validasi Peran

Klaster	Keanggotaan		Peran Pengguna		Label	Jumlah	%
	>=0.3	<0.3	>=0.3	<0.3	Akhir	Salah	Salah
C1	182	22	AK	AK, UM	AK	3	1.47%
C2	0	755		PC, PN, TL	PN, PC	247	32.72%
C6	117	195	TL	TL	TL	1	0.32%
C9	0	698		PN, SA	PN, SA	92	13.18%
C10	122	89	KR, PM	KR, PM	KR	32	15.17%
C11	342	45	AK, KB, KM, KR, PM, UM	AK, KR, PM	KR	10	2.58%
C12	740	65	KR, PM, PO	KR, PM	KR	19	2.36%
C14	213	217	UM, DM	AK, PM, TL	AK, PM, TL	63	14.65%
C15	382	79	TL	TL		0	0%

Pada pengukuran tingkat kesalahan pelabelan, nilai diberikan atas persentase anggota klaster yang memiliki kodifikasi berbeda dengan label klaster tersebut, misal: pengguna dengan kodifikasi peran PN pada klaster yang berlabel TL. Pada penelitian ini tingkat kesalahan pelabelan dari pengklasteran pendekatan 1 dan 2 dibandingkan dengan pengklasteran menggunakan metode DBSCAN, KMeans++ dan FCM. Sedangkan Pada pengukuran tingkat kemiripan menu hak akses pengguna, dilakukan perhitungan persentase menu hak akses yang dimiliki pada klaster tersebut. Sebagai contoh pada Tabel 3.5 untuk klaster C11 yang memiliki banyak variasi peran. Proses pengklasteran lanjutan dilakukan pada sub klaster C11 dengan nilai keanggotan <0.3. Kemudian pelabelan peran dilakukan disemua sub klaster yang menghasilkan peran KR. Terlihat ada 10 pengguna peran KR yang memiliki peran dari nama pengguna tidak sesuai sehingga tingkat kesalahan adalah 2.58%.

BAB 4 HASIL DAN PEMBAHASAN

4.1 Hasil Pembuatan Fitur dengan Matriks Pengguna

Terdapat dua matriks fitur pengguna yang dihasilkan dalam penelitian ini. Matriks pertama berukuran 4380x192 untuk mengakomodir kode cabang dan posisi pengguna. Matriks kedua berukuran 4380x190 karena tidak memasukkan kode cabang dan posisi pengguna sebagai fitur. Selanjutnya, atas matriks pertama dilakukan transformasi fitur menjadi menggunakan transformasi t-SNE sehingga diperoleh matriks berukuran 4380x3. Cuplikan hasil ketiga matriks ini disajikan pada LAMPIRAN 2. HASIL MATRIKS FITUR PENGGUNA.

4.2 Hasil Pengklasteran dengan Pendekatan 1

Penelitian ini seperti yang disebutkan sebelumnya menggunakan 4380 pengguna yang merupakan catatan dalam matriks akses pengguna CBS dari Bank XYZ. Beberapa paket Python seperti sklearn (scikit-learn.org) digunakan untuk transformasi t-SNE, pengelompokan KMeans ++ dan DBSCAN bersama dengan skfuzzy (pythonhosted.org/scikit-fuzzy) untuk pengelompokan FCM dalam penelitian ini.

Jumlah klaster pada FCM ditetapkan sejumlah 20 klaster berdasarkan pertimbangan bahwa nilai FPC mencapai nilai maksimal sedikit diatas 0,62 (Gambar 4.1) untuk jumlah peran yang hampir sama (Tabel 3.3). Sebagai perbandingan, juga dilakukan identifikasi peran melalui pengelompokan DBSCAN dan KMeans ++. Adapun metode yang diusulkan, pertama adalah metode kombinasi FCM dengan DBSCAN (Pendekatan 1). Hasil pengklasteran FCM dikelompokkan kembali menggunakan DBSCAN, untuk kemudian dilabeli ulang. Hasil pengklasteran menggunakan FCM ditunjukkan padaTabel 3.4, sedangkan hasil akhir pengklasteran menggunakan Pendekatan 1 ditunjukkan pada Tabel 4.1. Sedangkan Pendekatan 2 adalah FCM bertingkat dengan hasil pengklasteran FCM pertama dikelompokkan kembali menggunakan FCM. Untuk memvalidasi kinerja masing-masing metode, ditetapkan label klaster berdasarkan peran yang diekstrak

dari nama pengguna di setiap kluster. Tabel 4.1 menunjukkan bahwa Pendekatan 1 memiliki kinerja yang lebih baik dibanding pengklasteran tunggal tanpa modifikasi. Hasil dari tabel tersebut masih diproses lebih lanjut dengan DBSCAN pada beberapa klaster yang termasuk kategori klaster campuran, seperti C1, C2, C3 dan 17 klaster lainnya. Semua klaster campuran tersebut diproses ulang dengan hasil ditunjukkan pada Tabel 4.2. Sebagai contoh klaster C1 di Tabel 4.1 akan memiliki sub klaster C1' dan C2' di Tabel 4.2.

Tabel 4.1 Hasil pengklasteran proses FCM pertama di Pendekatan 1

Klaster	Label	Klaster Campuran?	Keterangan
C1	SA	Ya	Sub klaster ada di C1’ dan C2’
C2	KI	Ya	Sub klaster ada di C3' dan C4'
C3	KM	Ya	Sub klaster ada di C5', C6', dan C7'
C4	TL	Tidak	Diganti menjadi C8’ dan digabung dengan sub klaster lain
C5	AK	Tidak	Diganti menjadi C9’
C6	KR	Ya	Sub klaster digabung ke C5’, dan sebagian menjadi C10’
C7	TL	Tidak	Sub klaster digabung ke C8’
C8	PN	Tidak	Diganti menjadi C11’ dan digabung dengan sub klaster lain
C9	PC	Ya	Sub klaster ada di C12' dan C13', sebagian digabung ke C14’
C10	SA	Ya	Sub klaster digabung ke C1’ dan sebagian ke C8’ serta C15’
C11	TL	Ya	Diganti menjadi C8’ dan digabung dengan sub klaster C16’
C12	UM	Ya	Sub klaster ada di C17', C18', C19'
C13	KR	Ya	Sub klaster ada di C20' dan C21', digabung dengan sub klaster lain
C14	UM	Ya	Sub klaster ada di C22' dan C23', digabung dengan sub klaster lain
C15	TL	Tidak	Diganti menjadi C8’ dan digabung dengan sub klaster lain
C16	PM	Ya	Sub klaster digabung ke C5’ serta C24’ dan C25’
C17	PN	Tidak	Digabung dengan C11’
C18	PN	Ya	Sub klaster digabung ke C1’
C19	KR	Ya	Sub klaster digabung ke C5’ dan C26’
C20	KR	Tidak	Diganti menjadi C27’

Tabel 4.2 Hasil akhir pengklasteran di Pendekatan 1

Klaster	Label	Klaster Awal		Jumlah salah label (total = 607)			
C1'	SA	C1, C10, C18	C1_1, C10_2, C18_2	12	1	3	
C2'	KR	C1	C1_2	10			
C3'	KI	C2	C2_1	9			
C4'	PC	C2	C2_2	28			
C5'	KM	C3, C6, C16	C3_1, C6_2, C16_3	10	0	5	
C6'	KR	C3, C19	C3_2, C19_2	24	2		
C7'	TL	C3, C19	C3_3, C19_3	0	0		
C8'	TL	C4, C7, C10, C15	C4, C7, C10_3, C15	0	2	0	0
C9'	AK	C5	C5	7			
C10'	KR	C6, C12	C6_1, C12_4	101	1		
C11'	PN	C8, C17, C18	C8, C17, C18_3	0	0	0	
C12'	PC	C9	C9_1	5			
C13'	PN	C9	C9_2	39			
C14'	TL	C9, C11, C14	C9_3, C11_2, C14_3	0	0	0	
C15'	AK	C10	C10_1	0			
C16'	UM	C11	C11_1	1			
C17'	UM	C12	C12_1	1			
C18'	AK	C12	C12_2	97			
C19'	TL	C12	C12_3	0			
C20'	KR	C13, C18	C13_1, C18_1	101	7		
C21'	PM	C13, C19	C13_2, C19_4	6	29		
C22'	UM	C14	C14_1	69			
C23'	LN	C14	C14_2	3			
C24'	PM	C16	C16_1	17			
C25'	AK	C16	C16_2	3			
C26'	UM	C19	C19_1	1			
C27'	KR	C20	C20	13			

Pada Tabel 4.2**.** ditunjukkan hasil pengklasteran dengan Pendekatan 1 setelah beberapa klaster campuran diproses lebih lanjut dengan DBSCAN dan kemudian digabung. Syarat penggabungan dilakukan berdasarkan jarak klaster berdasarkan nilai tengah (*centroid*). Sebagai contoh C27’ adalah klaster dengan anggota awal dari C20.

Tabel 4.3 Perbandingan Hasil Pengklasteran Peran Pengguna di Bank XYZ

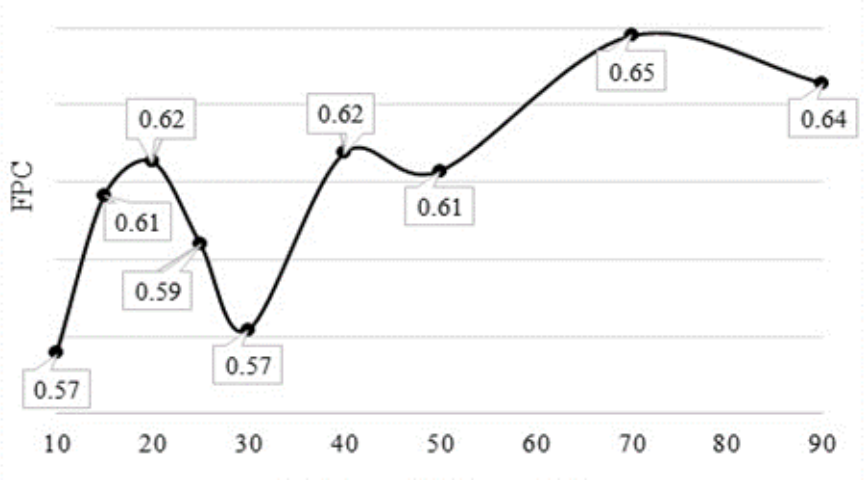
Algoritma	Hasil	% salah
DBScan (tanpa set jumlah klaster)	35 klaster dengan sebagian besar memiliki banyak peran. Label klaster ditentukan dari peran dengan prosentase terbesar.	41.62%
KMeans++ (jumlah klaster K = 13)	4 klaster dengan peran yang bervariasi dan 9 klaster memiliki label PC, SA, PM, PN, KI, KR, AK, TL, UM ditentukan dari peran dengan prosentase terbesar.	38.58%
FCM (jumlah klaster C = 13), nilai ambang keanggotaan = 0.15	9 klaster dengan peran yang bervariasi dengan label dari peran dengan prosentase terbesar untuk 2 klaster SA, 3 klaster TL, 3 klaster KR, 2 klaster PN, dan 1 klaster AK.	29.87%
Pendekatan 1 (FCM + DBSCAN + auto label)	27 klaster setelah proses klaster lanjutan dan penggabungan peran untuk AK, KI, KM, KR, LN, PC, PM, PN, SA, TL, dan UM.	13.85%

Prosentase kesalahan dihitung dari data yang memiliki label peran tidak sesuai. Diasumsikan peran pengguna yang benar diambil dari bagian di nama pengguna (Tabel 4.2). Kemudian proses selanjutnya adalah penentuan pengklasteran lanjutan dengan DBSCAN serta pemberian auto-label berdasarkan peran terbanyak di suatu klaster. Sebagai contoh terdapat klaster C1 hasil FCM dengan variasi peran AK(2.4%), CS(0.6%), HK(1.2%), KR(12.7%), PM(1.8%), PN(6.6%), SA(74.1%), TL(0.6%) sehingga perlu pengklasteran lanjutan DBSCAN. Klaster hasil FCM tidak memiliki peran yang valid karena prosentase semua peran < 75%.

Hasil FCM+DBSCAN pada klaster C1 adalah empat sub klaster C1_0 sampai C1_3. Auto-label berdasarkan peran terbanyak pada sub klaster C1_0 adalah KR dan seterusnya. Pada klaster C1 dengan auto-label akan memiliki data salah label sebanyak 22 sehingga prosentase kesalahan 22/166 = 13.25%. Informasi centroid dari klaster (hasil FCM) atau sub klaster (hasil FCM+DBSCAN) akan disimpan untuk proses penggabungan klaster jika jarak antar centroid kurang dari suatu nilai ambang dan memiliki label peran sama. Jika beberapa sub klaster (hasil

FCM+DBSCAN) dalam suatu klaster (hasil FCM) memiliki auto-label yang sama maka dilakukan penggabungan. Sebagai contoh sub klaster C1_1, C1_2 dan C1_3 akan digabung sehingga C1 memiliki sub klaster C1_A dengan peran KR dan C1_B dengan peran SA. Sebagai catatan, setelah analisis manual ditetapkan nilai ambang jarak untuk penggabungan adalah <50.

Terlihat 4 pengguna di sub klaster C1_0 memiliki label KR meski seharusnya memiliki label AK. Hal yang sama terjadi di sub klaster C1_0 untuk 2 pengguna dengan peran HK, 3 pengguna dengan peran PM dan 1 pengguna dengan peran TL akan memiliki auto-label KR.



Gambar 4.1 Evaluasi Fuzzy Partition Coefficient (FPC) (sumbu-y) dengan variasi jumlah klaster dari Fuzzy CMeans (FCM) (sumbu-x)

4.3 Validasi Pengklasteran Peran pada Pendekatan 1

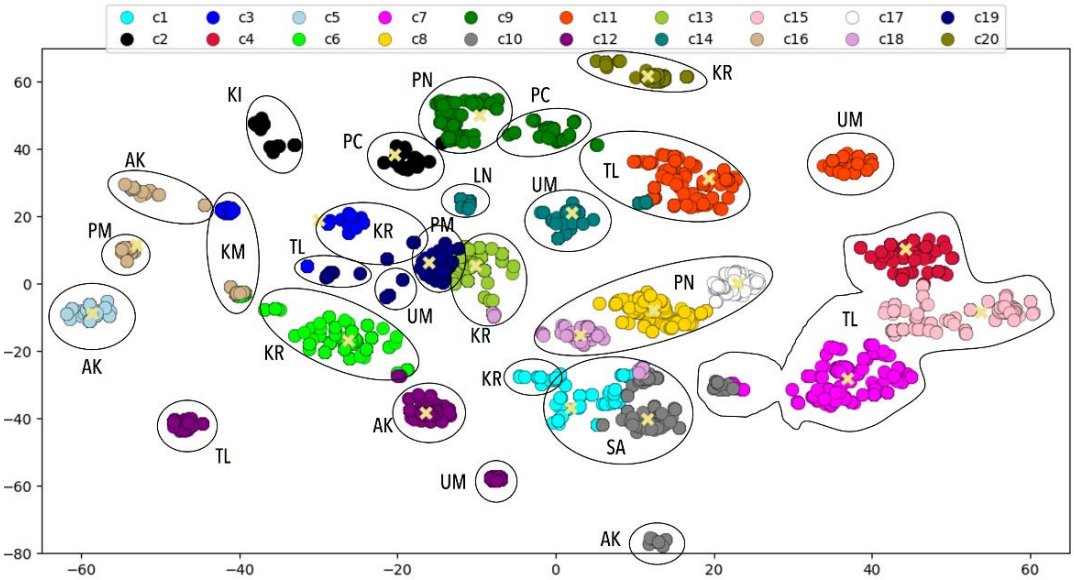
Gambar 4.2 menunjukkan hasil pengklasteran dengan variasi nilai parameter awal jumlah klaster untuk keperluan validasi secara visual. Indikator hasil klaster yang baik menggunakan Fuzzy Partition Coefficient (FPC) (Ross, 2010) sehingga nilai C=20 ditentukan untuk proses selanjutnya.

Validasi lebih lanjut dilakukan dengan membandingkan rekomendasi hak akses menu yang dihasilkan sesuai dengan pengklasteran FCM + DBSCAN. Hasil

FCM menunjukkan 20 klaster (Gambar 4.2) dengan persimpangan seperti C6-C12, C6-C16, C2-C9, C11-C14, C1-C10-C18, dan area lainnya. Hasil klaster FCM campuran setelah pelabelan adalah C1, C2, C3, C6, C9, C10, C12, C13, C14, C16, C18, dan C19. Klaster tersebut sesuai dengan dua syarat.

- Pertama, setidaknya satu peran memiliki proporsi pengguna 70% -80% dan peran lain yang berbeda dari 10% di atas.
- Kedua, tidak ada peran dengan proporsi pengguna lebih dari 60% tetapi setidaknya ada dua peran dengan proporsi 20% di atas.

Setelah DBSCAN pengklasteran kemudian diikuti proses penggabungan. Ada dua peran yang tidak disebutkan dalam visualisasi: DJ dan WP. Peran DJ hanya memiliki beberapa pengguna (Gambar 4.2) dibandingkan dengan yang lain sehingga mudah ditimpa. Sedangkan peran WP memiliki jumlah pengguna yang moderat dan alasan penggantian dijelaskan berikutnya.



Gambar 4.2 Visualisasi Pengklasteran Peran di Bank XYZ

Dari Gambar 4.2, pengguna peran PM tersebar ke beberapa kelompok yaitu C16 dan C19. Pemeriksaan akses pengguna PM mengungkapkan dua peran berbeda. Peran PM di C16 mewakili peran pemasaran atas produk dan layanan

4.4 Hasil Pengklasteran dengan Pendekatan 2

Pendekatan 2 adalah pengklasteran FCM pada data pengguna tanpa transformasi t-SNE menggunakan 189 fitur menu. Hasil FCM dengan nilai C=20 ditunjukkan di Tabel 4.4 dengan sembilan klaster yang membutuhkan pengklasteran FCM lanjutan sehingga disebut FCM bertingkat. Berdasarkan analisis awal ditetapkan nilai ambang keanggotaan 0.3 untuk menentukan pengklasteran FCM lanjutan. Terdapat 7 klaster yang memiliki data dengan nilai keanggotaan terbesar < 0.3 meski juga memiliki sebagian dengan nilai keanggotaan terbesar >=0.3. Proses FCM bertingkat juga dilakukan pada 2 klaster karena sama sekali tidak memiliki data dengan nilai keanggotaan terbesar >= 0.3. Sebagai catatan, terdapat beberapa klaster yang dianggap outlier karena hanya memiliki <10 anggota dengan nilai keanggotaan rendah (< 0.3). FCM lanjutan dilakukan dengan C=5. Terdapat dua klaster yang menjadi perhatian karena memiliki peran TL namun tersebar di C6 dan C15.

Tabel 4.4 Hasil Pengklasteran FCM pada Pengguna dengan Fitur Menu

Klaster	**Nilai Keanggotaan**		**Peran Pengguna**	
	>=0.3	**<0.3**	**>=0.3**	**<0.3**
C1	182	22	AK	AK, UM
C2	0	755		PC, PN, TL
C3	Outlier karena hanya terdapat < 10 data dan nilai keanggotaan <0.3			
C4	Outlier karena hanya terdapat < 10 data dan nilai keanggotaan <0.3			
C5	Tidak memiliki data			
C6	117	195	TL	TL
C7	Tidak memiliki data			
C8	Tidak memiliki data			
C9	0	698		PN, SA
C10	122	89	KR, PM	KR, PM
C11	342	45	AK, KB, KM, KR, PM, UM	AK, KR, PM
C12	740	65	KR, PM, PO	KR, PM
C13	Tidak memiliki data			
C14	213	217	UM, DM	AK, PM, TL
C15	382	79	TL	TL
C16	Outlier karena hanya terdapat < 10 data dan nilai keanggotaan <0.3			
C17	Outlier karena hanya terdapat < 10 data dan nilai keanggotaan <0.3			
C18	Tidak memiliki data			
C19	Tidak memiliki data			
C20	Tidak memiliki data			

Deskripsi hasil 7 klaster C1, C2, C6, C9, C10, C11, C12, C14, C15 dengan nilai keanggotaan < 0.3 ditunjukkan pada Tabel 4.4. Setiap data memiliki lima nilai keanggotaan namun hanya dipetakan ke satu sub klaster dengan nilai tertinggi. Terlihat bahwa ada beberapa sub klaster memiliki peran dominan lebih dari satu. Sebagai contoh hasil FCM lanjutan pada klaster C2, yaitu sub klaster C2_1 memiliki tiga peran KI (57 pengguna), PC (92 pengguna), dan WP (38 pengguna). Oleh karena itu perlu dilakukan analisis menu untuk setiap peran dominan per sub klaster. Analisis menu dilakukan dengan menghitung prosentase pengguna yang memiliki setiap menu di suatu sub klaster. Pada kasus tertentu dengan pengguna yang memiliki peran homogen akan dikenali set menu (*frequent itemsets*). Sebagai contoh AK di C1 dengan keanggotaan >=0.3 memiliki kombinasi menu GL3, GL1C dengan nilai dukung 0.995 atau terdapat 21 pengguna dengan set menu tersebut.

Tabel 4.5 Hasil Pengklasteran FCM Lanjutan dengan Keanggotaan Rendah

Klaster	**Hasil FCM Lanjutan (nilai C = 5)**
C1	Sebagian besar pengguna (17) menjadi anggota sub klaster dengan peran AK. Set menu dalam klaster terlihat di Tabel 4.6.
C2	Meski terdapat beberapa pengguna dengan peran berbeda namun peran dominan pada sub klaster adalah KI, PC, PN, TL dan WP (Gambar 4.3). C2_1 memiliki peran dominan KI, PC, WP C2_3 memiliki peran dominan PN C2_4 memiliki peran dominan PC_PN
C6	Hanya ada dua sub klaster dengan peran dominan TL meski akan terlihat set menu yang berbeda (Gambar 4.4)
C9	Terdapat lebih banyak variasi peran dibanding klaster C2 namun hanya ada dua peran dominan yaitu PN dan SA (Gambar 4.5)
C10	Jumlah pengguna dalam klaster lebih sedikit (89 pengguna) dibanding C9, sehingga dengan variasi peran serta dua peran dominan KR dan PM hanya memiliki pengguna < 30. Set menu dalam klaster terlihat di Tabel 4.6.
C11	Hanya terdapat empat peran KR AK PM UM dengan jumlah pengguna <50. Tidak ada peran yang dominan karena setiap sub klaster memiliki pengguna < 30. Set menu dalam klaster terlihat di Tabel 4.6.
C12	Hanya ada dua peran KR dan PM namun tidak ada peran yang dominan karena setiap sub klaster memiliki pengguna < 30. Set menu dalam klaster terlihat di Tabel 4.6.
C14	Meski terdapat lebih banyak jumlah pengguna, namun variasi peran banyak ditemukan dengan AK, PM dan TL menjadi peran dominan. Akan tetapi set menu dalam klaster terlihat di Tabel 4.6**.** tidak terlalu bervariasi dibanding klaster lain (C2, C6 dan C9).
C15	Hanya ada satu peran TL dalam klaster dengan tiga sub klaster namun semuanya memiliki menu yang sering muncul yaitu MPNT, PBBT, SBL1, SBL5, SBL6, SKNT, SNK1, SNK5, SNK6, TLKT

Menu-menu dalam klaster maupun sub klaster perlu dikenali untuk menghasilkan rekomendasi kebijakan standarisasi menu hak akses. Beberapa menu atau set menu yang sering muncul ditunjukkan di Tabel 4.6, Gambar 4.3, Gambar 4.4, dan Gambar 4.5. Sebagai contoh untuk peran TL yang tersebar di banyak klaster minimal harus memiliki akses menu PBBT (Pembayaran PBB Teller) dan TLKT (Host To Host Teller). Namun layanan kredit oleh Teller di kantor cabang baik perorangan (SNK1, SNK5, SNK6) maupun kelompok (SBL1, SBL5, SBL6) akan diwakili dengan menu-menu standard selain pelaporan transaksi tunai (CTRI, CTRL), transaksi tabungan giro (TDBJ, TGRT) dan pengolahan data pelanggan (KYCI, KYCL, KYCT). Namun terdapat peran yang memiliki menu dengan nilai *support* 100% sehingga semua pengguna mempunyai hak akses seperti informasi layanan informasi pelanggan (CSI) di peran AK, KB, KM, KR, PM, UM klaster C11.

Tabel 4.6 Set Menu dari Hasil Pengklasteran FCM Bertingkat

Klaster	Keanggotaan	Peran Dominan	Set Menu (frequent itemsets) dan prosentase minimal data dukung
C1	>=0.3	AK	GL3, GL1C (0.995)
	<0.3	AK	(1.00)GL1C, (0.21)KYCT, (0.16)L2M2, (0.16)L2PO, (0.16)L2T2
C6	Semua (data dengan keanggotaan >=0.3 dan <0.3)	TL	CTRI, CTRL, KYCI, KYCL, KYCT, TDBJ, TGRT Menu-menu berikut juga sering ditemui di pengguna klaster C15. PBBT, SBL1, SBL5, SBL6, SNK1, SNK5, SNK6, TLKT
C10	>=0.3	KR, PM	BGN2, BGN3, BGN4 (1.00)
C11	>=0.3	AK, KB, KM, KR, PM, UM	CSI (1.00) layanan informasi pelanggan
C12	>=0.3	KR, PM	SID1 (1.00) sistem informasi debitur
C14	>=0.3	UM	LUM (1.00) lap. bag. umum-personalia
	<0.3	PM	LN4 (1.00) laporan pinjaman
		AK	L2M2, L2PO, L2T2 (1.00) (0.88)MPNT
		TL	Menu-menu berikut juga sering ditemui di pengguna klaster C6 dan C15. (0.92)PBBT, (0.84)TLKT

Pengklasteran hak akses pengguna pada Bank XYZ dengan Pendekatan 1 (transformasi t-SNE dari 191 fitur yaitu menu serta status pengguna, FCM dilanjutkan DBSCAN serta penggabungan klaster) memberikan rata-rata salah pelabelan peran 13.85% (Tabel 4.7). Label peran diberikan berdasarkan jumlah peran terbanyak pengguna di suatu klaster. Untuk Pendekatan 2 (tanpa transformasi t-SNE, hanya 189 fitur menu, FCM bertingkat tanpa penggabungan klaster) memberikan rata-rata kesalahan 10.31% (Tabel 4.7). Pengguna dengan jumlah label peran kurang tepat terbanyak ditemukan pada klaster C2 (Tabel 4.8) yang tidak memiliki data dengan nilai keanggotaan tinggi (>=0.3). Meskipun terjadi pengurangan jumlah pengguna dengan salah label, namun proses pada Pendekatan 2 lebih sedikit jika dibandingkan Pendekatan 1. Selain itu Pendekatan 2 hanya menggunakan informasi menu hak akses pengguna tanpa status pengguna di kantor cabang maupun posisi sehingga mengurangi asumsi peran ganda akibat kebutuhan operasional.

Oleh karena itu berdasarkan uraian poin-poin berikut, maka ditunjukkan hasil dari Pendekatan 2 lebih baik dari Pendekatan 1 berdasarkan Tabel 4.3, Tabel 4.7, dan Tabel 4.8:

a. Langkah pada Pendekatan 2 lebih sedikit dibanding Pendekatan 1, sehingga perulangan yang dibutuhkan untuk cek klaster serta penggabungan lebih sedikit.
b. Jumlah klaster akhir serta jumlah data salah label pada Pendekatan 2 (13 klaster, 452 salah label) lebih sedikit dibanding dengan Pendekatan 1 (27 klaster, 607 salah label).

Tabel 4.7 Perbandingan Hasil Pengklasteran Pendekatan-1 dan Pendekatan-2

Algoritma	Langkah	% salah
Pendekatan 1	1. FCM 2. Auto label 3. Cek klaster campuran 4. DBSCAN 5. Auto label 6. Penggabungan klaster-subklaster	13.85%
Pendekatan 2	1. FCM 2. Cek keanggotaan 3. FCM 4. Auto Label 5. Penggabungan subklaster	10.31%

Tabel 4.8 Hasil akhir pengklasteran di Pendekatan 2

Klaster	Keanggotaan		Peran Pengguna		Label	# Salah
	>=0.3	<0.3	>=0.3	<0.3	Akhir	(452)
C1	182	22	AK	AK, UM	AK	3
C2	0	755		PC, PN, TL	PN, PC	247
C6	117	195	TL	TL	TL	1
C9	0	698		PN, SA	PN, SA	92
C10	122	89	KR, PM	KR, PM	KR	17
C11	342	45	AK, KB, KM, KR, PM, UM	AK, KR, PM	KR	10
C12	740	65	KR, PM, PO	KR, PM	KR	19
C14	213	217	UM, DM	AK, PM, TL	AK, PM, TL	63
C15	382	79	TL	TL	TL	0

4.5 Standarisasi Menu Hak Akses dengan Pendekatan 2

Berdasarkan analisis hasil pengklasteran dengan Pendekatan 2 yang menggunakan FCM bertingkat tanpa transformasi t-SNE pada data fitur, maka dihasilkan rekomendasi standarisasi menu hak akses untuk beberapa peran terbanyak dalam sistem perbankan XYZ (Tabel 4.13). Pengklasteran hak akses pengguna dengan Pendekatan 1 dan Pendekatan 2 menghasilkan rekomendasi untuk kelompok peran: AK, PM, KR, KM, TL, KI, UM, PC, PN, LN, SA. Sebagai contoh hak akses untuk kelompok peran AK (Tabel 4.9).

Dari kemiripan sebaran hak akses menu antara klaster C5 dengan sub klaster C16_1, maka menu GL1C (transaksi general ledger cabang) dan GL3 (laporan general ledger) dapat diberikan untuk pegawai Bank XYZ dengan peran staf Akutansi. Hak akses untuk sub klaster C12_1 dapat diberikan untuk pegawai pada cabang pembantu dengan peran penyelia akuntansi. Sedangkan hak akses untuk sub klaster C10_1 diberikan pada pegawai cabang dengan peran penyelia akutansi. Rekomendasi serupa oleh Pendekatan 2 diberikan pada Tabel 4.6.

BAB 5 KESIMPULAN

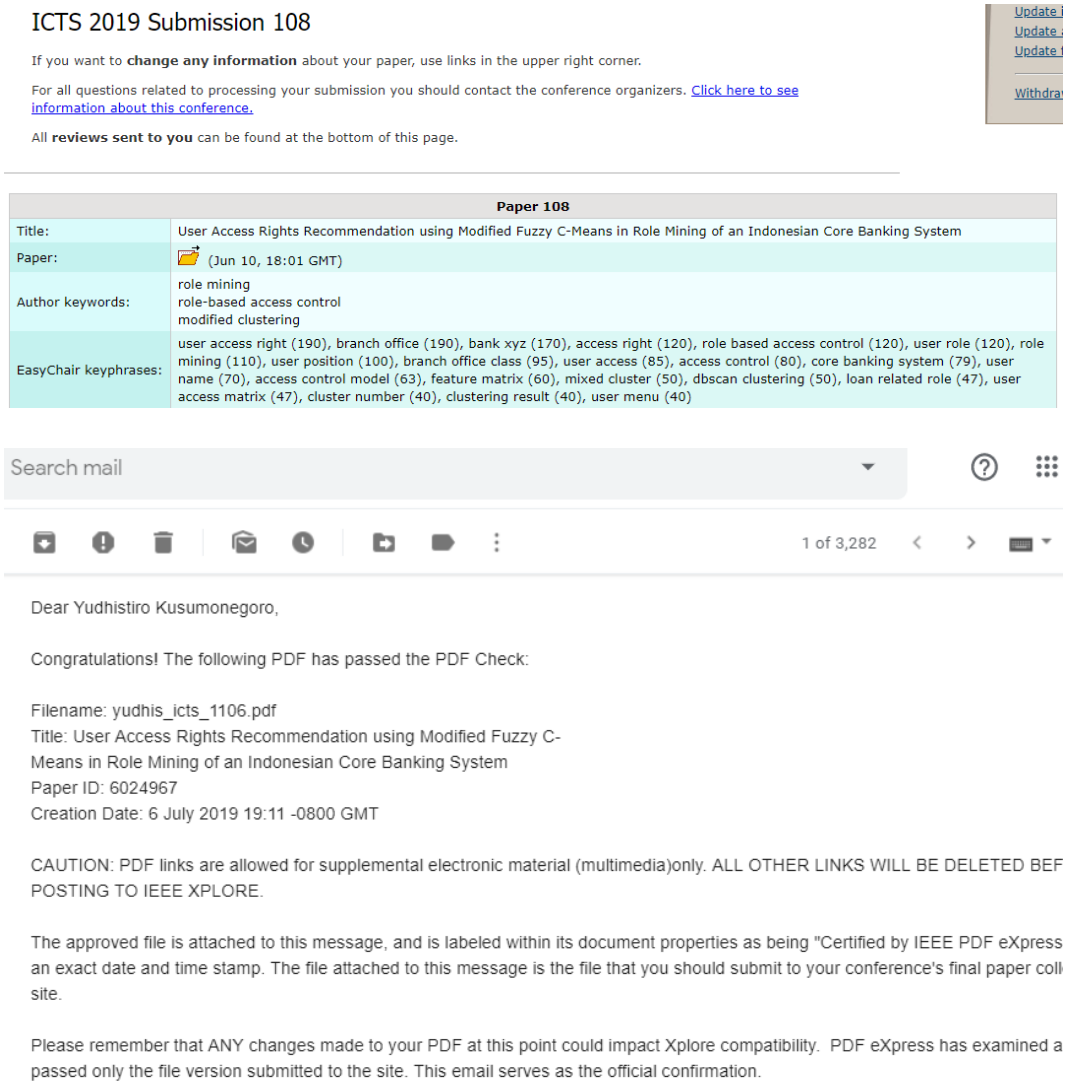
Pada penelitian ini modifikasi Fuzzy C-Means dilakukan untuk rekomendasi hak akses pengguna Core Banking System di bank XYZ. Beberapa hal penting yang dapat disimpulkan adalah sebagai berikut.

1. Rekomendasi standarisasi hak akses dapat dilakukan dengan pengklasteran FCM bertingkat pada fitur menu tanpa informasi status pengguna (tingkat kesalahan 10.31%) (Pendekatan 2).
2. Pengklasteran dengan transformasi t-SNE pada fitur menu dan status pengguna memberikan hasil visual untuk membantu validasi (Pendekatan 1).
3. Salah klaster sering terjadi pada pengguna dengan peran ganda karena status pengguna menjadi fitur dalam pengklasteran (Pendekatan 1).
4. Pengklasteran berdasarkan menu saja membuat beberapa pengguna dengan peran ganda menjadi data outlier dengan nilai keanggotaan rendah (Pendekatan 2).

Penelitian saat ini memiliki asumsi bahwa peran ganda perlu dikenali dan pasti ada sehingga proses pengklasteran dilakukan beberapa tahap baik di Pendekatan 1 dan Pendekatan 2. Pada penelitian selanjutnya dapat dilakukan pengecekan menu di suatu klaster tanpa mempertimbangkan label peran. Sehingga jika menu dalam suatu klaster sudah homogen maka tidak perlu dilakukan proses pengklasteran ulang.

LAMPIRAN 1. PUBLIKASI MAKALAH ICTS

Sebagai syarat kelulusan Program Magister Manajemen Teknologi, Institut Teknologi Sepuluh Nopember (ITS), sebagian hasil dari penelitian ini yaitu pengklasteran hak akses dengan transformasi t-SNE atau disebut Pendekatan 1 telah dikirimkan ke International Conference on Information & Communication Technology and System (ICTS) 2019. Seminar akan dilaksanakan pada tanggal 18 Juli 2018 di Departemen Informatika ITS.



ICTS 2019 Submission 108

If you want to **change any information** about your paper, use links in the upper right corner.

For all questions related to processing your submission you should contact the conference organizers. Click here to see information about this conference.

All **reviews sent to you** can be found at the bottom of this page.

Update i
Update
Update
Withdra

Paper 108	
Title:	User Access Rights Recommendation using Modified Fuzzy C-Means in Role Mining of an Indonesian Core Banking System
Paper:	(Jun 10, 18:01 GMT)
Author keywords:	role mining role-based access control modified clustering
EasyChair keyphrases:	user access right (190), branch office (190), bank xyz (170), access right (120), role based access control (120), user role (120), role mining (110), user position (100), branch office class (95), user access (85), access control (80), core banking system (79), user name (70), access control model (63), feature matrix (60), mixed cluster (50), dbscan clustering (50), loan related role (47), user access matrix (47), cluster number (40), clustering result (40), user menu (40)

Search mail

1 of 3,282

Dear Yudhistiro Kusumonegoro,

Congratulations! The following PDF has passed the PDF Check:

Filename: yudhis_icts_1106.pdf
Title: User Access Rights Recommendation using Modified Fuzzy C-Means in Role Mining of an Indonesian Core Banking System
Paper ID: 6024967
Creation Date: 6 July 2019 19:11 -0800 GMT

CAUTION: PDF links are allowed for supplemental electronic material (multimedia)only. ALL OTHER LINKS WILL BE DELETED BEF POSTING TO IEEE XPLORE.

The approved file is attached to this message, and is labeled within its document properties as being "Certified by IEEE PDF eXpress an exact date and time stamp. The file attached to this message is the file that you should submit to your conference's final paper coll site.

Please remember that ANY changes made to your PDF at this point could impact Xplore compatibility. PDF eXpress has examined a passed only the file version submitted to the site. This email serves as the official confirmation.

DAFTAR PUSTAKA

Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining Association Rules Between Sets of Items in Large Databases. *SIGMOD Rec.*, *22*(2), 207–216. https://doi.org/10.1145/170036.170072

Bezdek, J. C. (1981). *Pattern Recognition with Fuzzy Objective Function Algorithms*. Norwell, MA, USA: Kluwer Academic Publishers.

Du, X., & Chang, X. (2014). Performance of AI algorithms for mining meaningful roles. In *2014 IEEE Congress on Evolutionary Computation (CEC)* (pp. 2070–2076). https://doi.org/10.1109/CEC.2014.6900321

Dunn, J. C. (1973). A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *Journal of Cybernetics*, *3*(3), 32–57. https://doi.org/10.1080/01969727308546046

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A Density-based Algorithm for Discovering Clusters a Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining* (pp. 226–231). AAAI Press. Retrieved from http://dl.acm.org/citation.cfm?id=3001460.3001507

Gavrila, S. I., & Barkley, J. F. (1998). Formal Specification for Role Based Access Control User/Role and Role/Role Relationship Management. In *Proceedings of the Third ACM Workshop on Role-based Access Control* (pp. 81–90). New York, NY, USA: ACM. https://doi.org/10.1145/286884.286902

Knox, S. W. (2018). Machine Learning: A concise introduction. John Wiley & Sons.

Kuhlmann, M., Shohat, D., & Schimpf, G. (2003). Role Mining - Revealing Business Roles for Security Administration Using Data Mining Technology.

In *Proceedings of the Eighth ACM Symposium on Access Control Models and Technologies* (pp. 179–186). New York, NY, USA: ACM. https://doi.org/10.1145/775412.775435

Li, R., Li, H., Wang, W., Ma, X., & Gu, X. (2013). RMiner: A Tool Set for Role Mining. In *Proceedings of the 18th ACM Symposium on Access Control Models and Technologies* (pp. 193–196). New York, NY, USA: ACM. https://doi.org/10.1145/2462410.2462431

Lin, D., Rao, P., Bertino, E., & Lobo, J. (2007). An Approach to Evaluate Policy Similarity. In *Proceedings of the 12th ACM Symposium on Access Control Models and Technologies* (pp. 1–10). New York, NY, USA: ACM. https://doi.org/10.1145/1266840.1266842

Ma, X., Li, R., & Lu, Z. (2010). Role Mining Based on Weights. In *Proceedings of the 15th ACM Symposium on Access Control Models and Technologies* (pp. 65–74). New York, NY, USA: ACM. https://doi.org/10.1145/1809842.1809854

Ma, X., Li, R., Lu, Z., & Wang, W. (2012). Mining constraints in role-based access control. *Mathematical and Computer Modelling*, *55*(1), 87–96. https://doi.org/10.1016/j.mcm.2011.01.053

Mitra, B., Sural, S., Vaidya, J., & Atluri, V. (2016). A Survey of Role Mining. *ACM Comput. Surv.*, *48*(4), 50:1--50:37. https://doi.org/10.1145/2871148

Molloy, I., Chen, H., Li, T., Wang, Q., Li, N., Bertino, E., … Lobo, J. (2008). Mining Roles with Semantic Meanings. In *Proceedings of the 13th ACM Symposium on Access Control Models and Technologies* (pp. 21–30). New York, NY, USA: ACM. https://doi.org/10.1145/1377836.1377840

Park, J. S., & Giordano, J. (2006). Role-based profile analysis for scalable and accurate insider-anomaly detection. In *2006 IEEE International Performance Computing and Communications Conference* (pp. 7 pp. – 470). https://doi.org/10.1109/.2006.1629440

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., …

Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Ross, T. J. (2010). *Fuzzy Logic With Engineering Applications, 3rd ed.* Wiley.

Saenko, I., & Kotenko, I. (2017). Administrating Role-based Access Control by Genetic Algorithms. In *Proceedings of the Genetic and Evolutionary Computation Conference Companion* (pp. 1463–1470). New York, NY, USA: ACM. https://doi.org/10.1145/3067695.3082509

Schaad, A., Moffett, J., & Jacob, J. (2001). The Role-based Access Control System of a European Bank: A Case Study and Discussion. In *Proceedings of the Sixth ACM Symposium on Access Control Models and Technologies* (pp. 3–9). New York, NY, USA: ACM. https://doi.org/10.1145/373256.373257

Vaidya, J., Atluri, V., & Guo, Q. (2007). The Role Mining Problem: Finding a Minimal Descriptive Set of Roles. In *Proceedings of the 12th ACM Symposium on Access Control Models and Technologies* (pp. 175–184). New York, NY, USA: ACM. https://doi.org/10.1145/1266840.1266870

Vaidya, J., Atluri, V., & Guo, Q. (2010). The Role Mining Problem: A Formal Perspective. *ACM Trans. Inf. Syst. Secur.*, *13*(3), 27:1--27:31. https://doi.org/10.1145/1805974.1805983

Vaidya, J., Atluri, V., & Warner, J. (2006). RoleMiner: Mining Roles Using Subset Enumeration. In *Proceedings of the 13th ACM Conference on Computer and Communications Security* (pp. 144–153). New York, NY, USA: ACM. https://doi.org/10.1145/1180405.1180424

van der Maaten, L., & Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research*, *9*, 2579–2605. Retrieved from http://www.jmlr.org/papers/v9/vandermaaten08a.html

Zhang, D., Ramamohanarao, K., & Ebringer, T. (2007). Role Engineering Using Graph Optimisation. In *Proceedings of the 12th ACM Symposium on Access Control Models and Technologies* (pp. 139–144). New York, NY, USA: ACM. https://doi.org/10.1145/1266840.1266862

BIODATA

Yudhistiro Trah Kusumonegoro. Penulis dilahirkan di Surabaya, 18 September 1978. Menamatkan kuliah S1 di Jurusan Teknik Kimia Institut Teknologi Sepuluh Nopember Surabaya, penulis memilih berkarir di bidang teknologi informasi, khususnya di bidang jaringan, infrastruktur dan keamanan informasi. Kemudian penulis melanjutkan studi di S2 Magister Manajemen Teknologi, Institut Teknologi Sepuluh Nopember Surabaya.

Email: yudhistiro@gmail.com