

Pemetaan Profil Mahasiswa Untuk Peningkatan Strategi Promosi Perguruan Tinggi Menggunakan *Predictive Apriori*

Irfan Pratama¹, Putri Taqwa Prasetyaningrum²

Program Studi Sistem Informasi

Universitas Mercu Buana Yogyakarta

Yogyakarta, Indonesia

e-mail: ¹irfanp@mercubuana-yogya.ac.id, ²putri@mercubuana-yogya.ac.id

Diajukan: 8 Februari 2021; Direvisi: 28 Maret 2021; Diterima: 31 Maret 2021

Abstrak

Program promosi melalui berbagai media untuk meningkatkan dan menguatkan rekrutmen mahasiswa yang terdistribusi pada setiap program studi secara ideal dan setara diperlukan untuk membuat perguruan tinggi selalu bertumbuh. Demografis mahasiswa dapat menjadi karakteristik khusus dalam memilih program studi. Begitu pun latar belakang pendidikan mahasiswa dan keluarga. Berdasarkan masalah yang telah dikemukakan sebelumnya, maka peneliti tertarik untuk melakukan penelitian terkait data pendaftaran mahasiswa. Secara spesifik pengolahan terhadap data secara konseptual dapat dilakukan menggunakan proses Data Mining. Salah satu proses dari Data Mining yang dapat memberikan pengetahuan tentang keterkaitan antar variabel data adalah Association-Rule Mining. Metode yang digunakan yaitu Predictive Apriori sebagai metode asosiasi untuk mengekstrak profil mahasiswa dari universitas untuk mendapatkan wawasan yang lebih baik tentang apa yang harus dilakukan tim pemasaran untuk memperkuat kampanye pemasaran. Data tersebut akan diolah sehingga dapat dikenali karakteristik atau pola dari data-data yang ada dengan harapan dapat memberikan pengetahuan untuk dapat meningkatkan efektivitas promosi dan rekrutmen mahasiswa. Metode tersebut menghasilkan total 44456 rules, yang selanjutnya dilakukan filter berdasarkan rules yang mengandung atribut Prodi dan nilai akurasi minimum 90% (0.9). Rules yang dihasilkan akan dijadikan saran ajuan terhadap bagian pemasaran untuk melakukan pola roadshow atau promosi sesuai pola yang telah terbentuk.

Kata kunci: Association rule, Data mining, Pembelajaran mesin, Predictive Apriori.

Abstract

As a growing University, there is definitely a need for promotional programs through various media to increase and strengthen the recruitment of students who are distributed in each course of study ideally and equally. The demographic aspects of the students may be a special characteristic in choosing a course of study. Based on the problems that have been raised before, the researchers are interested in conducting research related to student registration data. Specifically, the processing of data conceptually can be done using the Data Mining process. One of the processes of Data Mining that can provide knowledge about the interrelationship between data variables is Association-Rule Mining. Predictive Apriori used as an association method to extract student profiles from universities to gain better insight into what marketing teams should do to strengthen marketing campaigns. The results of this study are 44456 number of rules produced by the method, and then filtered by the containment of Prodi attribute and the minimum accuracy score of 90% (0.9). The rules that found and filtered then proposed to the marketing department to follow as the promotion pattern on their roadshow schedule.

Keywords: Association rule, Data mining, machine learning, Predictive Apriori.

1. Pendahuluan

Menurut data yang dikeluarkan FORLAP RISTEKDIKI, Indonesia memiliki lebih dari 3000 perguruan tinggi dan 585 di antaranya merupakan perguruan tinggi swasta dan masih banyak lagi jenis perguruan tinggi swasta lainnya seperti institut, akademi, dan politeknik. 127 di antaranya berada di D.I. Yogyakarta, tentunya lebih dari itu jika memasukkan PTN ke dalam perhitungan [1]. Dengan banyaknya institusi tersebut maka persaingan di antara mereka tentunya cukup tinggi. Artinya kerja tim *marketing* khususnya untuk PTS harus lebih keras agar bisa menarik dan membujuk calon mahasiswa untuk

bergabung. Selain dari akreditasi lembaga, efektivitas dan efisiensi tim pemasaran dapat memberikan pengaruh yang sangat besar terhadap penerimaan calon mahasiswa. Jadi, mereka butuh semacam strategi agar bisa unggul dari kompetitor lain dalam hal promosi atau kampanye pemasaran. Promosi atau kampanye pemasaran merupakan salah satu hal yang harus dilakukan untuk membawa kesuksesan bagi perusahaan atau institusi. Menurut Kotler dan Keller, Konsep pemasaran menekankan bahwa kunci untuk mencapai tujuan organisasi yang dinyatakan adalah bahwa perusahaan harus lebih efektif daripada pesaing dalam menciptakan, menyampaikan, dan mengkomunikasikan nilai pelanggan ke pasar sasaran yang dipilih [2]. Dengan penggunaan informasi yang tepat, cara promosi yang efektif atau ke mana tujuan kampanye pemasaran dapat ditentukan untuk mendapatkan hasil yang lebih baik.

Beberapa tahun belakangan terdapat suatu bidang ilmu yang sangat tren di dunia teknologi informasi yang terkait dengan bidang-bidang lain termasuk bisnis yaitu *Data Science*. *Data Science* adalah sebuah bidang ilmu yang secara khusus mempelajari data, khususnya data yang bersifat kuantitatif atau numerik baik yang terstruktur maupun tak terstruktur sehingga data-data tersebut dapat memberikan suatu pemahaman terkait masalah atau fakta yang ada [3][4]. Secara spesifik pengolahan terhadap data secara konseptual dapat dilakukan menggunakan proses Data Mining. Salah satu proses dari Data Mining yang dapat memberikan pengetahuan tentang keterkaitan antar variabel data adalah *Association-Rule Mining*. *Association-Rule* adalah sebuah pembelajaran mesin berbasis aturan untuk menemukan keterkaitan antar variabel data pada sebuah basis data atau kumpulan data. Pengetahuan yang dihasilkan dari proses ini adalah kumpulan-kumpulan aturan yang menyatakan hubungan sebab akibat antar variabel data [5][6].

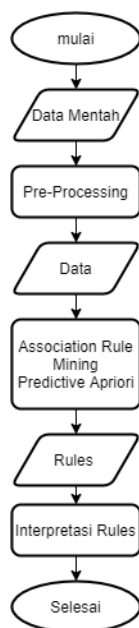
Ada beberapa metode yang digunakan untuk penambangan aturan asosiasi yaitu Apriori, FP-Growth, dan *Predictive Apriori*. Apriori akan baik digunakan jika ada hubungan antara atribut yang dianalisis. Dan salah satu penggunaan aturan asosiasi adalah untuk menganalisis perilaku pelanggan dan strategi promosi dan pemasaran produk menggunakan metode apriori [7][8][9][10]. Bentuk implementasi lain dari aturan asosiasi adalah untuk mendukung strategi promosi perguruan tinggi menggunakan metode Apriori [11]. Penelitian tentang analisis pola penjualan produk juga menggunakan apriori sebagai metode dan menghasilkan 6 aturan terbaik dengan menggunakan dukungan minimum 40% dan kepercayaan minimum 60% [10]. Penelitian [12] dilakukan dengan membandingkan waktu evaluasi antara metode apriori dan FP-growth terhadap penjualan paket sembako. Hasil penelitian menyatakan bahwa dengan metode apriori, apriori menghasilkan 8 aturan terbaik dengan dukungan minimum 0,06 dan kepercayaan minimum 0,01 dengan tingkat akurasi 35%. Sedangkan metode FP-Growth menghasilkan 14 aturan terbaik dengan menggunakan minimal *support* 0,06 dengan tingkat akurasi 84%.

Universitas Mercu Buana Yogyakarta memiliki 13 program studi pada jenjang sarjana (S1) dengan bidang keilmuannya masing-masing. Sebagai Universitas yang berkembang, pasti perlu adanya program promosi melalui berbagai media untuk meningkatkan dan menguatkan rekrutmen mahasiswa yang terdistribusi pada setiap program studi secara ideal dan setara. Daerah asal mahasiswa pelamar dimungkinkan dapat menjadi karakteristik khusus dalam memilih program studi. Begitu pun profil dari mahasiswa itu sendiri seperti asal jurusan, asal sekolah dan lain-lain. Sehubungan dengan hal tersebut, maka untuk melakukan penguatan rekrutmen mahasiswa pada program studi tertentu dapat dilakukan dengan mengetahui karakteristik spesifik dari data mahasiswa sehingga dapat diharapkan hasil promosi yang tepat sasaran. Di mana pada sebelumnya kecenderungan promosi yang dilakukan masih sebatas mendatangi kota – kota dengan frekuensi jumlah mahasiswa yang besar tanpa spesifik melihat keterkaitan antara minat calon mahasiswa terhadap Prodi, sehingga proses promosinya masih umum.

Oleh karena itu, penelitian ini akan dilakukan dengan memanfaatkan metode asosiasi untuk mengekstrak pola dari profil mahasiswa untuk mendapatkan wawasan yang lebih baik tentang apa yang harus dilakukan tim pemasaran untuk memperkuat kampanye pemasaran berikutnya. Data tersebut akan diolah sedemikian rupa sehingga dapat dikenali karakteristik atau pola dari data-data yang ada dengan harapan dapat memberikan pengetahuan untuk dapat meningkatkan efektivitas promosi dan rekrutmen mahasiswa. Metode *Predictive Apriori* dipilih karena metode tersebut merupakan bentuk yang lebih mutakhir dari metode Apriori. Diharapkan dengan mengimplementasikan metode yang lebih mutakhir maka hasil dari metode tersebut akan lebih baik.

2. Metode Penelitian

Pada tahap ini akan dijelaskan alur dari jalannya penelitian ini. Dimulai dengan pengambilan data atau *Data Acquisition*, kemudian tahap *Pre-Processing* untuk memastikan data siap digunakan, kemudian data yang telah bersih dan siap digunakan akan di analisis menggunakan metode *Predictive Apriori* untuk mendapatkan aturan asosiasi berdasarkan dari data yang dimiliki, dan yang terakhir adalah interpretasi dari hasil analisis sesuai kebutuhan. Diagram alir dari jalannya penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Flowchart penelitian.

2.1. Data

Data yang digunakan pada penelitian ini adalah data mahasiswa dari seluruh program studi sarjana (S1) yang ada di Universitas Mercu Buana Yogyakarta dari rentang tahun 2016 – 2020 dengan total data berjumlah 10.146 baris data. Tabel 1 berikut adalah gambaran umum dari bentuk data yang digunakan pada penelitian ini.

Tabel 1. Sampel data.

No	Prodi	Provinsi	Pendapatan orang Tua	Pekerjaan Orang Tua	Asal Sekolah
1	Ilmu Komunikasi	JAMBI	<1000000	Buruh	SMA Payakumbuh
2	Informatika	JAMBI	2000000 - 2500000	Lainnya	MA Nurul Hidayah
3	Ilmu Komunikasi	JAMBI	>3000000	Swasta	SMA N 2 Tanjung Jabung
4	Psikologi	GORONTALO	>3000000	PNS	MA Mubarak
5	Psikologi	JAMBI	1000000 - 2000000	PNS	SMA N 1 Tanjung jabung

2.2. Pre-Processing

Pada tahap *Pre-Processing* ini terdapat beberapa langkah yang dilakukan untuk mempersiapkan data mentah menjadi data yang siap digunakan untuk proses *data mining* selanjutnya. Langkah-langkah tersebut meliputi: pengisian *missing values*, dan transformasi data.

2.2.1. Transformasi Data

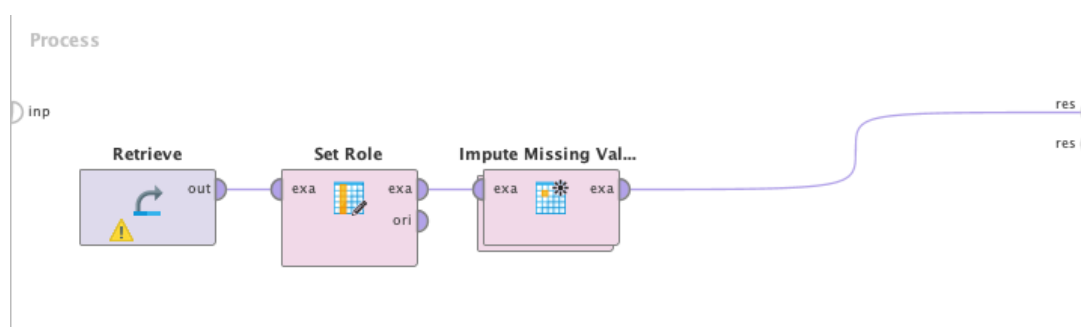
Pada tahap transformasi data ini, beberapa kolom data (atribut) akan ditransformasikan menjadi bentuk yang lebih *general* dan mudah untuk diproses. Bentuk transformasi yang akan dilakukan terhadap data mentah yang dimiliki adalah pada kolom Pendapatan Orang Tua dan Asal Sekolah. Seperti yang terlihat pada Tabel 1, data Asal sekolah masih berupa rincian asal sekolah dari masing-masing mahasiswa. Sehingga akan dikelompokkan menjadi 3 jenis sekolah yaitu SMA, SMK, dan MA. Begitu juga dengan kolom Pendapatan Orang Tua yang akan di transformasikan menjadi kode seperti <1000000 = A, 100000 – 2000000 = B, 2000000 – 2500000 = C, 2500000 – 3000000 = D, >3000000 = E. Tabel 2 berikut ini adalah contoh dari hasil transformasi *dataset*.

Tabel 2. Hasil transformasi data.

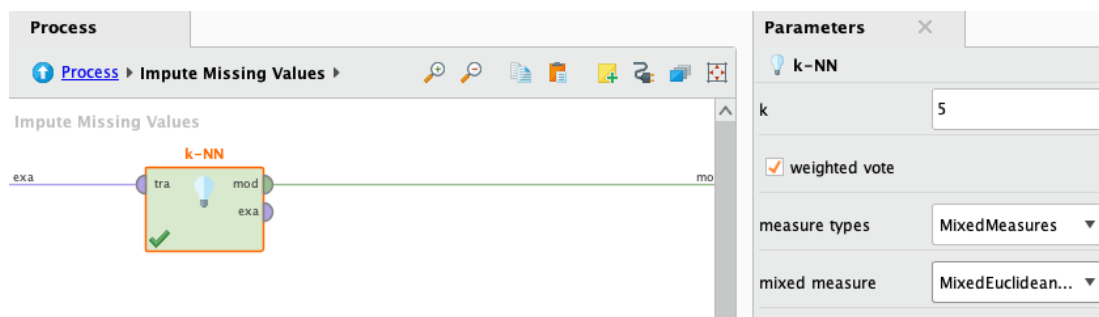
No	Prodi	Provinsi	Pendapatan orang Tua	Pekerjaan Orang Tua	Asal Sekolah
1	Ilmu Komunikasi	JAMBI	A	Buruh	SMA
2	Informatika	JAMBI	C	Lainnya	MA
3	Ilmu Komunikasi	JAMBI	E	Swasta	SMA
4	Psikologi	GORONTALO	E	PNS	MA
5	Psikologi	JAMBI	B	PNS	SMA

2.2.2. Pengisian Missing Values

Dataset yang digunakan memiliki *missing values* pada bagian asal sekolah, pendapatan orang tua dan pekerjaan orang tua, karena sifat *form* pada asal sekolah saat pengisian data adalah *text box*. Sehingga dimungkinkan adanya ketidaklengkapan data yang diberikan. Sedangkan pada data pendapatan orang tua dan pekerjaan orang tua tidak di set sebagai *required field* sehingga juga diperbolehkan kosong. Agar data dapat digunakan dengan utuh tanpa ada yang dihapus karena mengandung *missing values*, dibutuhkan mekanisme untuk memperkirakan nilai untuk menggantikan *missing values* yang ada pada data tersebut. Terdapat berbagai mekanisme yang dapat dilakukan untuk menangani *missing values* yang terjadi pada sebuah *dataset*. Dari mekanisme yang sangat sederhana seperti menghapus baris data yang mengandung *missing values* yang akan mengakibatkan data berkurang sehingga informasi-informasi nya yang dapat di akuisisi juga berkurang. Terdapat juga sebuah mekanisme pengisian *missing values* menggunakan *mean* (untuk data numerik) atau *mode* (untuk data nominal) yang tidak efektif untuk jumlah *missing values* yang banyak karena akan menimbulkan bias. Hingga mekanisme mutakhir yang menggunakan *machine learning* [13]. Mekanisme pengisian *missing values* yang digunakan pada penelitian ini adalah menggunakan metode K-NN [14][15]. Menggunakan *tools* RapidMiner dalam pengisian *missing values*, skema penanganan *missing values* menggunakan metode K-NN dapat dilihat pada Gambar 2 dan Gambar 3 berikut.



Gambar 2. Skema *missing values imputation*.



Gambar 3. Metode K-NN.

Pada Gambar 2 dapat dilihat skema pengisian *missing values* yang mengharuskan *dataset* yang akan diisi *missing values*-nya menggunakan metode K-NN untuk di set labelnya terlebih dahulu, karena penelitian ini ingin menonjolkan *profiling* mahasiswa untuk setiap Prodi nya, sehingga label dari datanya akan ditujukan kepada data Prodi. Pada Gambar 3 dapat dilihat pengaturan parameter untuk metode K-NN menggunakan $k = 5$ dengan parameter lainnya bersifat *default*. Hasil perubahan dari pengisian *missing values*-nya dapat dilihat pada Gambar 4 dan Gambar 5 yang menunjukkan data sebelum dan pengisian *missing values*.

Row No. ↓	PROV	PEKERJAAN	PENDAPAT...	ASAL SEKO...	PRODI
8461	JAWA TENGAH	Lainnya	1	SMA	Psikologi
8460	PAPUA BARAT	?	1	SMA	Psikologi
8459	JAWA TIMUR	Swasta	1	SMA	Akuntansi
8458	PAPUA	Swasta	3	SMK	Manajemen
8457	SUMATERA ...	PNS	4	SMA	Manajemen
8456	DI YOGYAKA...	PNS	3	SMK	Psikologi
8455	PAPUA BARAT	PNS	2	SMK	Informatika
8454	PAPUA BARAT	PNS	4	SMK	Ilmu Komuni...
8453	SUMATERA ...	Lainnya	1	SMA	Psikologi
8452	LAMPUNG	Swasta	1	MA	Psikologi
8451	JAWA TENGAH	?	1	SMA	Psikologi
8450	DI YOGYAKA...	Buruh	0	SMK	Ilmu Komuni...

Gambar 4. sebelum penanganan *missing values*

Row No. ↓	PROV	PEKERJAAN	PENDAPAT...	ASAL SEKO...	PRODI
8461	JAWA TENGAH	Lainnya	1	SMA	Psikologi
8460	PAPUA BARAT	Swasta	1	SMA	Psikologi
8459	JAWA TIMUR	Swasta	1	SMA	Akuntansi
8458	PAPUA	Swasta	3	SMK	Manajemen
8457	SUMATERA ...	PNS	4	SMA	Manajemen
8456	DI YOGYAKA...	PNS	3	SMK	Psikologi
8455	PAPUA BARAT	PNS	2	SMK	Informatika
8454	PAPUA BARAT	PNS	4	SMK	Ilmu Komuni...
8453	SUMATERA ...	Lainnya	1	SMA	Psikologi
8452	LAMPUNG	Swasta	1	MA	Psikologi
8451	JAWA TENGAH	Lainnya	1	SMA	Psikologi
8450	DI YOGYAKA...	Buruh	0	SMK	Ilmu Komuni...
8449	SUMATERA ...	Lainnya	3	SMA	Ilmu Komuni...
8448	JAWA TENGAH	Swasta	1	SMK	Peternakan

Gambar 5. Setelah Penanganan *missing values* menggunakan K-NN *imputation*.

2.3. Predictive Apriori

Predictive Apriori dikembangkan oleh Scheffer [16]. Metode ini menggunakan nilai *support* yang lebih besar dan ditukar dengan nilai *confidence* yang lebih tinggi pula dan akurasi dihitung dalam *Bayesian Network*. Hasil dari algoritma ini memaksimalkan nilai akurasi untuk data-data yang akan datang. Algoritma ini menghasilkan jumlah *rules* sesuai dengan yang di spesifikasikan oleh pengguna. Algoritma ini didefinisikan sebagai berikut:

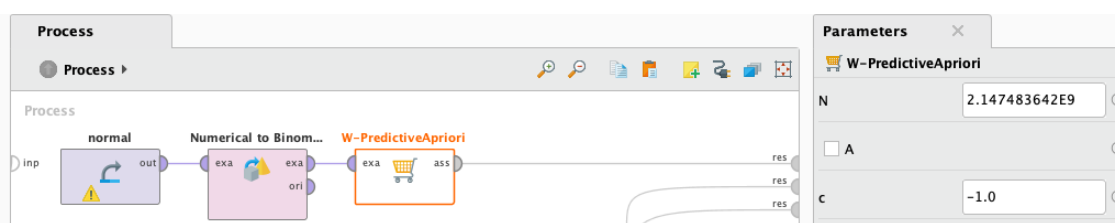
Misalkan D adalah *database* di mana r adalah *record* individual yang dihasilkan oleh proses statis P , misalkan $X \rightarrow Y$ menjadi aturan asosiasi. Akurasi prediktif $c(X \rightarrow Y) = Pr(r \text{ memenuhi } Y | r \text{ memenuhi } X)$ adalah probabilitas bersyarat dari $Y \subseteq r$ mengingat bahwa $X \subseteq r$ ketika distribusi r diatur oleh P [16].

Fungsi perhitungan akurasi prediktif adalah sebagai berikut:

$$E(c(r) | \hat{c}(r), s(X)) = \frac{\int_{B[c,s(X)]} (\hat{c}(r)) P(c) dc}{\int_{B[c,s(X)]} (\hat{c}(r)) P(c) dc} \tag{1}$$

Di mana $E(c(r) | \hat{c}(r), s(X))$ adalah akurasi prediktif yang diharapkan dari *rule* $r: X \rightarrow Y$, nilai *confidence* dinotasikan sebagai \hat{c} dan nilai *support* dari *rule* nya di notasikan sebagai $s(X)$.

Pada penelitian ini jumlah *rules* yang diinginkan adalah pada nilai maksimal dari yang bisa dihasilkan oleh data. Dengan menggunakan RapidMiner, skema *Predictive Apriori* dapat dilihat pada Gambar 5.



Gambar 6. Predictive Apriori.

Pada Gambar 5 dapat dilihat parameter *Predictive Apriori* pada jumlah *rules* diset pada nilai *default* dari yang diberikan oleh RapidMiner yaitu 2.147.483.642, meskipun jumlah *rules* yang dihasilkan mungkin saja tidak akan sebanyak itu karena tergantung dari yang ditemukan berdasarkan data yang dimiliki.

2.4. Interpretasi Rules

Setelah proses pembentukan *rules* asosiasi menggunakan algoritma *Predictive Apriori*, *rules* yang dihasilkan akan diinterpretasikan sesuai dengan kebutuhan dari permasalahan yaitu pemetaan profil mahasiswa berdasarkan masing-masing Prodi, sehingga kombinasi-kombinasi yang tidak sesuai kebutuhan akan dihilangkan dari *rule set*.

3. Hasil dan Pembahasan

Dari total *rules* yang telah dihasilkan oleh metode *Predictive Apriori* yang berjumlah 44456 *rules*, tidak semua dari *rules* tersebut yang akan digunakan untuk diterapkan dan disampaikan kepada pengambil keputusan strategis atau dalam hal ini tim pemasaran dan juga nilai akurasi dari setiap *rules*-nya harus yang di atas 90% (0.9). Proses interpretasi dari *rules* yang didapatkan harus disesuaikan dengan hasil yang diharapkan untuk digunakan selanjutnya. Total terdapat 85 *rules* yang disaring dari total *rules* yang dihasilkan oleh metode *Predictive Apriori*. Contoh sampel dari *rules* hasil penyaringan dapat dilihat pada Tabel 3 berikut ini.

Tabel 3. Hasil penyaringan *rules*.

No	Rules
1	JAWA TENGAH=true Swasta=true SMA=false Manajemen=true c=true 20 ==> SMK=9+63 true 20 acc:(0.95438)
2	NUSA TENGGARA TIMUR=true SMA=false Peternakan=true a=true 16 ==> SMK=true 16 acc:(0.94431)
3	DI YOGYAKARTA=true Buruh=true SMA=false Teknologi Hasil Pertanian=true 33 ==> SMK=true 32 acc:(0.94312)
4	DI YOGYAKARTA=true Buruh=true SMA=false Pendidikan Bahasa Inggris=true a=true 13 ==> SMK=true 13 acc:(0.93322)
5	DI YOGYAKARTA=true Buruh=true SMA=false Informatika=true 40 ==> SMK=true 39 acc:(0.9527)
6	SULAWESI SELATAN=true PNS=true Psikologi=true 12 ==> e=true 12 acc:(0.92847)
7	NUSA TENGGARA TIMUR=true Buruh=true SMK=true Akuntansi=true 11 ==> a=true 11 acc:(0.92298)
8	DKI JAKARTA=true SMA=false Ilmu Komunikasi=true 9 ==> SMK=true 9 acc:(0.90902)
9	SMK=true Sistem Informasi=true 79 ==> SMA=false 79 acc:(0.987)
...	...
85	DI YOGYAKARTA=true SMA=false Manajemen=true b=true 82 ==> SMK=true 75 acc:(0.90476)

Dari Tabel 3 dapat dilihat hasil dari penyaringan *rules-rules* yang relevan dengan pemecahan masalah yang ada pada penelitian ini dengan mengambil *rules-rules* yang mengandung *attribut* Prodi dengan nilai akurasi minimum 90% (0.9). dari keseluruhan hasil penyaringan *rules* yang dilakukan, ternyata belum mewakili *profiling* untuk setiap Prodi yang ada.

Contoh dari interpretasi *rules* yang dihasilkan dari metode *Predictive Apriori* adalah sebagai berikut:

Rule 1 : Jawa Tengah = true, Swasta = true, SMA = false, Manajemen = true, c = true, SMK = true.

Interpretasinya adalah =Mahasiswa yang memilih Prodi Manajemen Sebagian besar berasal dari SMK di Jawa Tengah yang berasal dari keluarga berlatar belakang pekerjaan Swasta dan kategori penghasilan C”

(di mana C merupakan kategori pendapatan yang ada di rentan 2.5 – 3 juta rupiah). Dari pengetahuan tersebut, langkah yang dapat dilakukan untuk meningkatkan efektivitas promosi perguruan tinggi adalah melakukan *roadshow* ke SMK yang berada di Jawa Tengah untuk memperkuat *intake* mahasiswa dari wilayah tersebut untuk Prodi Manajemen. Terkait dengan latar belakang ekonomi mahasiswa manajemen yang terpola berada di kategori C, maka tidak perlu dilakukan diskon pendaftaran atau beasiswa pendaftaran karena sudah dianggap mampu. Contoh lain dari interpretasi *rule* menjadi saran strategi promosi adalah *rules* 4: DI Yogyakarta = *true*, Buruh = *true*, Pendidikan Bahasa Inggris = *true*, a = *true*, SMK = *true*. Interpretasinya adalah = Mahasiswa Pendidikan Bahasa Inggris sebagian besar berasal dari Yogyakarta yang orang tuanya bekerja sebagai buruh dan kategori pendapatan berada pada kategori a (berada di *range* 0 – 1 juta rupiah). Sehingga strategi *marketing*-nya untuk dapat memperkuat animo calon mahasiswa untuk Prodi tersebut adalah memberikan diskon atau promo tertentu pada saat *roadshow* di Yogyakarta dan di sekolah SMK untuk siswa-siswa sekolah yang berasal dari keluarga yang ada di kategori ekonomi “a”.

Hasil dari penelitian ini nantinya akan menjadi usulan dan pengetahuan yang dapat digunakan untuk teknis strategi pemasaran tergantung pada profil mahasiswa pada setiap Prodi untuk memperkuat promosi perguruan tinggi. Profil mahasiswa yang dimaksud adalah demografi mahasiswa seperti asal daerah asal sekolah dan juga latar belakang ekonomi keluarga yang dilihat dari kategori pekerjaan orang tua dan kategori pendapatan orang tua.

4. Kesimpulan

Dari permasalahan yang dikemukakan sebelumnya bahwa pola dari profil mahasiswa dapat memudahkan dan mengoptimalkan kinerja dari tim *marketing* universitas dalam pemasaran dan promosi perguruan tinggi dapat diselesaikan menggunakan model *association rule mining*. Meskipun masih belum terpetakan secara adil untuk setiap Prodi dan sesuai yang diharapkan bahwa setiap Prodi akan memiliki profil yang khas dari sisi demografis mahasiswanya, karena terdapat *missing values* sehingga data yang real tidak dapat digunakan dan persebaran mahasiswa setiap Prodi yang jumlahnya bisa jadi tidak merata sehingga frekuensi dari setiap *itemset* tidak dapat di ekstrak sepenuhnya. Hasil ini akan diharapkan dapat membantu dalam strategi promosi perguruan tinggi ke depannya.

Saran untuk kelanjutan dari penelitian ini adalah dengan menggunakan jumlah *attribute* yang lebih banyak lagi dan mencoba model-model *data mining* gabungan untuk mengoptimalkan hasil seperti proses *balancing* data untuk masing-masing Prodi menggunakan *Class Balancing Methods*.

Daftar Pustaka

- [1] Pddikti, “Grafik Jumlah Perguruan Tinggi,” 2020. [Online]. Available: <https://forlap.ristekdikti.go.id/perguruantinggi/homegraphpt>. [Accessed: 10-Mar-2020].
- [2] P. Kotler And K. L. Keller, *Marketing Management*, 15th Ed. Pearson Education Inc., 2016.
- [3] V. Dhar, “Data Science And Prediction,” *Commun. Acm*, Vol. 56, No. 12, Pp. 64–73, 2013, Doi: 10.1145/2500499.
- [4] J. Leek, “The Key Word In ‘Data Science’ Is Not Data, It Is Science,” 2012. [Online]. Available: <https://simplystatistics.org/2013/12/12/the-key-word-in-data-science-is-not-data-it-is-science/>.
- [5] I. H. Witten, E. Frank, And M. A. Hall, *Data Mining: Practical Machine Learning Tools And Techniques*, 3rd Ed. San Francisco, Ca, Usa: Morgan Kaufmann Publishers Inc., 2011.
- [6] L. C. Yu, C. L. Chan, C. C. Lin, And I. C. Lin, “Mining Association Language Patterns Using A Distributional Semantic Model For Negative Life Event Classification,” *J. Biomed. Inform.*, Vol. 44, No. 4, Pp. 509–518, 2011, Doi: 10.1016/j.jbi.2011.01.006.
- [7] S. Yakub, “Analisis Data Mining Untuk Strategi Promosi Produk Kosmetik Di Wardah Kosmetik Menggunakan Metode Apriori,” Vol. 3, No. 1, Pp. 163–181, 2020.
- [8] U. Baetulloh, A. I. Gufroni, And R. -, “Penerapan Metode Association Rule Mining Pada Data Transaksi Penjualan Produk Kartu Perdana Kuota Internet Menggunakan Algoritma Apriori,” *Simetris J. Tek. Mesin, Elektro Dan Ilmu Komput.*, Vol. 10, No. 1, Pp. 173–188, 2019, Doi: 10.24176/Simet.V10i1.2890.
- [9] A. F. Afif, E. R. Swedia, And M. Cahyanti, “Implementasi Algoritma Association Rule Untuk Promosi Produk Berbasis Website Pada Bengkel Delta Jaya Motor,” *J. Ilm. Teknol. Dan Rekayasa*, Vol. 24, No. 2, Pp. 152–160, 2019, Doi: 10.35760/Tr.2019.V24i2.2034.
- [10] Y. Apridona, “Penerapan Data Mining Menggunakan Metode Assiciation Rule Dengan Algoritma Apriori Untuk Analisa Pola Penjualan Barang,” *Jurteks*, Vol. V, No. 2, Pp. 193–198, 2019.
- [11] H. Kusumo, E. Sedyono, And M. Marwata, “Analisis Algoritma Apriori Untuk Mendukung Strategi Promosi Perguruan Tinggi,” *Walisongo J. Inf. Technol.*, Vol. 1, No. 1, P. 49, 2019, Doi:

-
- 10.21580/Wjit.2019.1.1.4000.
- [12] H. Maulidiya, A. Jananto, G. Special, I. A. Bawang, M. Sedap, And M. Asosiasi, “Asosiasi Data Mining Menggunakan Algoritma Apriori Dan Fp-Growth Sebagai Dasar Pertimbangan Penentuan Paket Sembako,” Pp. 978–979, 2020.
- [13] R. J. A. Little And D. B. Rubin, *Statistical Analysis With Missing Data*. New York, Ny, Usa: John Wiley & Sons, Inc., 1986.
- [14] S. Zhang, “Nearest Neighbor Selection For Iteratively Knn Imputation,” *J. Syst. Softw.*, Vol. 85, No. 11, Pp. 2541–2552, 2012, Doi: 10.1016/J.Jss.2012.05.073.
- [15] T. Data Detective, “Preprocessing: Encode And Knn Impute All Categorical Features Fast,” 2019.
- [16] T. Scheffer, “Finding Association Rules That Trade Support Optimally Against Confidence,” *Intell. Data Anal.*, Vol. 9, No. 4, Pp. 381–395, 2005, Doi: 10.3233/Ida-2005-9405.