

Sistem Informasi Prediksi Mahasiswa Putus Kuliah Menggunakan Metode Data Mining Dengan Algoritma Chaid

Dhila Franzely Dhimas Putra¹, Yogasetya Suhandi², Miri Susanti³

^{1,2,3}Fakultas Teknologi, ITB Swadharma

Jl. Malaka No.3, RT.7/RW.3, Roa Malaka, Kec. Tambora, Kota Jakarta Barat,
dhila.amp@gmail.com, yoga@swadharma.ac.id, Mirisusanti@swadharma.ac.id

Abstract

College life cannot be separated from problems faced by students. This problem is one of the obstacles in completing the study. This problem is one of the obstacles in completing the lecture. The problem of students dropping out of college is a challenge for institutions that must be found a solution, because the dropout phenomenon causes significant losses for both students and institutions. The purpose of the research is to produce a prediction system for students who will drop out of college using data mining by processing data from ITB Swadharma for the period 1992 to 2019 and modeling using the Chaid algorithm, the results are visualized with the tableau application. The results of the study reveal the fact that the study time is the main factor that influences the prediction of dropout students with a chi-square value of 15,714, factor of accreditation status with a chi-square value of 16,874, gender is the factor with a chi-square value of 5.292.

Keyword: dropout, chaid algorithm, data mining

I. Pendahuluan

Mahasiswa merupakan bagian dari sivitas akademik pada sebuah perguruan tinggi, merupakan calon pemimpin bangsa di masa yang akan datang. Kurikulum merespon dan menyodorkan berbagai wawasan, cara berfikir baru mahasiswa diharuskan menyesuaikan diri dengan perubahan tuntutan belajar dari masa sebelumnya yaitu jenjang Pendidikan Sekolah Menengah Atas (SMA), menjadi siswa yang mandiri dalam segala hal aktivitas akademik baik dalam materi perkuliahan, tugas, laporan praktikum, tugas akhir sebagai syarat kelulusan dan ketepatan waktu kelulusan untuk menghindari drop out (DO).

Keberadaan Mahasiswa memiliki tempat tersendiri di lingkungan masyarakat, namun bukan berarti memisahkan diri dari masyarakat [1], kehidupan mahasiswa tidak terlepas dari berbagai macam permasalahan yang ada. Permasalahan tersebut dapat bersumber dari berbagai faktor seperti dalam diri sendiri, keluarga, pekerjaan, teman sepergaulan atau lingkungan sosial. Permasalahan ini dapat menyebabkan mahasiswa mengalami kendala dalam menyelesaikan kuliahnya sehingga beresiko menjadi mahasiswa putus kuliah (drop out).

Saat ini, perusahaan dan dunia ekonomi secara luas menganggap pengetahuan itu penting dan diterima sebagai keunggulan kompetitif perusahaan, maka setiap tahunnya perusahaan mengharapkan mahasiswa lulusan universitas mampu menjadi sumber daya manusia dalam organisasi mereka [2]. Tetapi data angka putus kuliah di Propinsi DKI Jakarta pada tahun 2017 mencapai 3.7%, dimana sebanyak 40.584 mahasiswa putus kuliah dari 1.084.123 mahasiswa yang aktif di perguruan tinggi negeri dan swasta yang berada di Propinsi DKI Jakarta [3]. Jumlah mahasiswa putus kuliah ini, selain merugikan secara individu, juga dapat merugikan institusi perguruan tinggi karena akan menurunkan indeks kinerja perguruan tinggi tersebut. Oleh karena itu, perlu dilakukan prediksi terhadap faktor - faktor yang mempengaruhi mahasiswa drop out sehingga dapat dijadikan informasi yang bermanfaat bagi perguruan tinggi.

Permasalahan tentang mahasiswa putus kuliah di perguruan tinggi menjadi tantangan bagi institusi yang harus ditemukan solusinya, karena fenomena putus kuliah menimbulkan kerugian signifikan baik bagi mahasiswa maupun institusi. Bagi mahasiswa putus kuliah akan mengganggu perkembangan psikososialnya seperti muncul rasa frustrasi akan pengembangan karier di masa depan maupun menurunnya academic self efficacy, sedangkan bagi institusi fenomena mahasiswa putus kuliah mengarah pada pemborosan biaya

berkaitan dengan re-alokasi dana mahasiswa yang mengikuti perkuliahan dan rekrutmen mahasiswa baru maupun menurunkan tingkat reputasi perguruan tinggi [4].

Proses ekstraksi atau penggalian data yang mana data tersebut belum diketahui sebelumnya, untuk menemukan pola yang menarik dan pengetahuan dari data database yang besar serta digunakan untuk membuat suatu keputusan bisnis yang sangat penting yang dikenal dengan Data Mining [5]. Penggunaan konsep data mining untuk memprediksi mahasiswa yang akan putus kuliah pernah dilakukan menggunakan algoritma C.45 yang mendapatkan penyebab tertinggi mahasiswa berhenti kuliah adalah pembayaran biaya kuliah yang tidak lancar dan memiliki nilai yang rendah [6]. Penelitian selanjutnya menerapkan metode data mining dengan menggunakan Data Mining Addins for MS Office 2007 menemukan bahwa faktor latar belakang pendidikan sewaktu SMA yaitu nilai NEM dan STTB berpengaruh terhadap kemampuan akademik mahasiswa yang dapat menyebabkan mahasiswa tersebut putus kuliah [7].

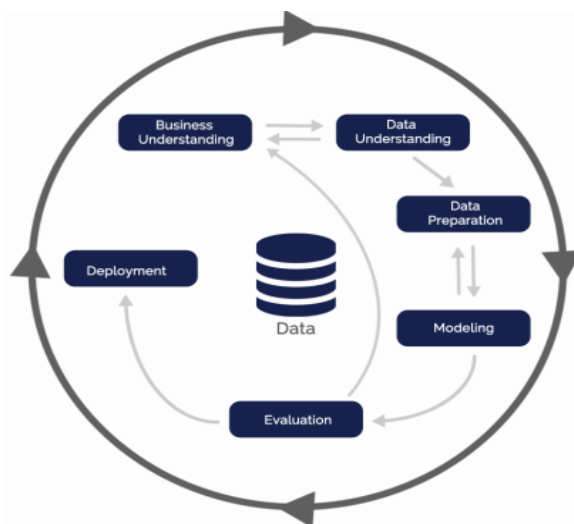
Metode CHAID (Chi-squared Automatic Interaction Detection) bekerja dengan mempelajari hubungan antara variabel dependen dengan beberapa variabel independen kemudian mengklasifikasi sampel data berdasarkan pada tingkat signifikansi statistik chi-square terhadap variabel dependennya [8]. Penggunaan Algoritma ini memungkinkan digunakan dalam tiga tahap, yaitu pemisahan (splitting), penggabungan (merging) dan penghentian (stopping). Penggambaran diagram pohon dimulai dari root node (node akar), melalui tiga tahap tersebut pada setiap node yang terbentuk dan secara berulang [9].

Oleh karena itu maka diperlukan suatu sistem informasi berbasis web [10] berisi data mahasiswa ITB Swadharma untuk memprediksi mahasiswa putus kuliah (Drop Out / DO) agar dapat dilakukan tindakan pencegahan untuk memperkecil jumlah mahasiswa putus kuliah. Tujuan penelitian ini adalah untuk menyediakan sistem informasi yang mampu memprediksi mahasiswa yang akan putus kuliah di Kampus ITB Swadharma. Hasil prediksi ini dapat digunakan ITB Swadharma untuk lebih mengenal kondisi mahasiswa dan dapat dijadikan sebagai pengetahuan dini dalam proses pengambilan keputusan untuk mengantisipasi mahasiswa putus kuliah.

II. Metodologi Penelitian

Metode yang diterapkan dalam penelitian ini adalah CRISP-DM. CRISP-DM (*Cross-Industry Standard Process Model for Data Mining*) adalah standar proses *data mining* sebagai strategi pemecahan masalah secara umum dari bisnis atau untuk penelitian, CRISP-DM menyediakan standar proses baku untuk *data mining* yang dapat diterapkan ke dalam strategi pemecahan masalah umum pada bisnis atau pada unit penelitian. CRISP-DM membandingkan metodologi *data mining* lain lebih lengkap dan terdokumentasi dengan baik. Setiap fase terstruktur dan terdefinisi dengan jelas sehingga mudah diaplikasikan bahkan bagi pemula sekalipun [11]. Metode ini menjelaskan tentang proses **data mining** dalam enam tahap yaitu (1) *Business Understanding*; (2) *Data Understanding*; (3) *Data Preparation*; (4) *Modeling*; (5) *Evaluation*; (6) *Deployment* [12].

Mariscal, Marban dan Fernandez [13] menyatakan CRISP-DM sebagai defacto menjadi standar untuk pengembangan proyek data mining dan knowledge discovery karena paling banyak digunakan dalam pengembangan data mining. Gambar 1 merupakan gambaran alur proses yang terjadi pada CRISP-DM.



Gambar 1. Alur CRISP DM

Penerapan metode ini dalam penelitian dapat dijabarkan sebagai berikut:

1. *Business Understanding*

Permasalahan dalam penelitian ini mengacu pada segmentasi mahasiswa, berdasarkan data yang terdapat pada sistem akademik yang ada untuk tindakan preventif mahasiswa putus kuliah di ITB Swadharma. Pada tahapan ini diperlukan pemahaman tentang proses yang berjalan agar dapat digunakan untuk mengetahui permasalahan yang ada dan merancang tindakan preventif yang akan diterapkan kepada data mahasiswa dengan tujuan untuk menurunkan angka mahasiswa putus kuliah di ITB Swadharma.

2. *Data Understanding*

Pemahaman data adalah tahapan mengumpulkan data awal dan mempelajari data tersebut untuk bisa mengenal dan memahami apa saja yang bisa dilakukan pada data-data tersebut. Pengumpulan data yang di butuhkan adalah data mahasiswa program studi Sistem Informasi, Teknik Informatika dan Manajemen Informatika pada periode tahun 1992 sampai dengan tahun 2019, yaitu berkaitan dengan data mahasiswa berupa nim, nama, alamat, jurusan kuliah, jurusan asal SMA, semester, tahun masuk, tahun lulus dengan kondisi null, dan status mahasiswa yang aktif.

3. *Data Preparation*

Data yang telah diolah ditahap sebelumnya kemudian dianalisa dan memilih variable yang sesuai dengan kebutuhan analisis. Setelah itu membersihkan data agar siap digunakan untuk pemodelan dengan menggunakan aplikasi SPSS Modeler. Pada tahap ini dilakukan penambahan data baru pada tabel `tb_dp_mhs_do`, penambahan tersebut dilakukan untuk mengetahui lama kuliah mahasiswa yaitu dengan cara menuliskan perintah di SPSS modeller yaitu `tahun_masuk – tahun_lulus`. Pada tahap ini juga dilakukan proses *merge* dengan menggabungkan tabel `mahasiswa`, `mahasiswa_2`, `tabel_nilai` dan `table_krs`, proses penggabungan tabel tersebut dilakukan dengan bantuan SPSS modeller serta menggunakan metode *inner join*.

4. *Modeling*

Setelah menyelesaikan proses *data preparation* kemudian dilanjutkan pada proses pembuatan *data modeling*. Tahapan untuk membuat model prediktif yang disebut *Data modeling* yaitu untuk memprediksi data dari mahasiswa yang putus kuliah pada ITB Swadharma. Tahapan ini diggunakan statistika dan *machine learning*, untuk mendapatkan *insight* yang berguna dari data yang ada untuk mencapai tujuan penelitian. Setelah dilakukan modeling dengan algoritma Chaid maka akan terbentuk *summary model* pada spss modeler.

5. *Evaluation*

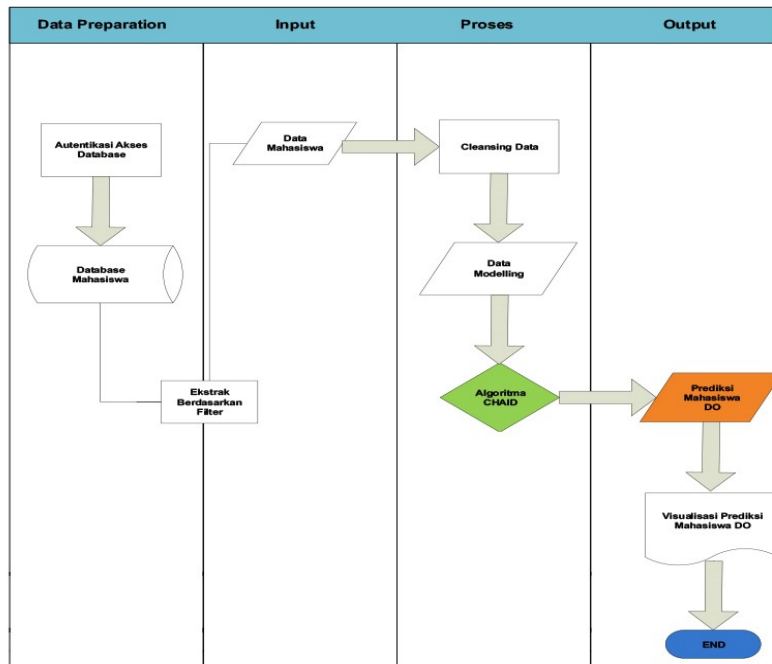
Dari perolehan data hasil proses modeling dan visualisasi didapatkan model penelitian ini berupa analisa dan perancangan yang mengarah ke pengambilan keputusan strategis. Hasil analisa diharapkan dapat dimanfaatkan dan di implementasikan oleh pihak akademik ITB Swadharma atau lembaga lain yang terkait dengan akademik

6. *Deployment*

Hasil evaluasi kemudian diterjemahkan kedalam aplikasi sistem informasi prediksi putus kuliah menggunakan IBM SPSS Modeler untuk modeling data, aplikasi Tableau untuk visualisasi data dan MySql untuk basis data

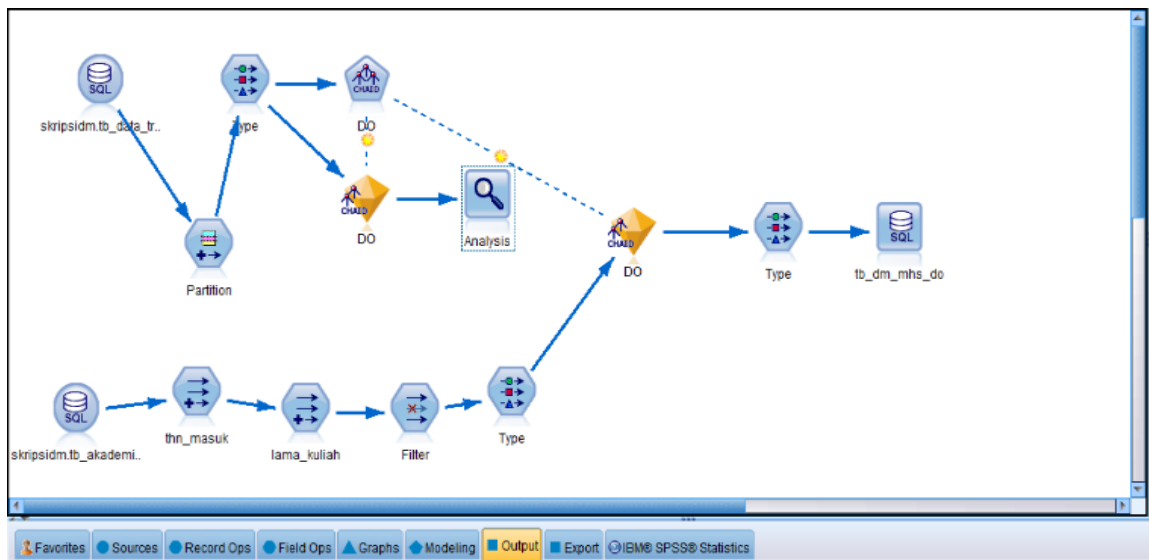
III. Hasil dan Pembahasan

Gambaran umum proses pengolahan data yang akan dijadikan sistem prediksi mahasiswa DO digambarkan menggunakan bentuk *Flowchart* pada gambar 2 :



Gambar 2. Flowchart Sistem

Setelah menyelesaikan proses *data preparation* kemudian dilanjutkan pada proses pembuatan *data modeling*, model ini dibuat untuk membuat model *prediktif*, yaitu untuk memprediksi mahasiswa yang putus kuliah di ITB Swadharma. Penggunaan statistika dan *machine learning* pada tahapan ini, untuk mendapatkan *insight* yang berguna dari data untuk mencapai tujuan penelitian. Tampilan *data modeling* seperti pada gambar 3:



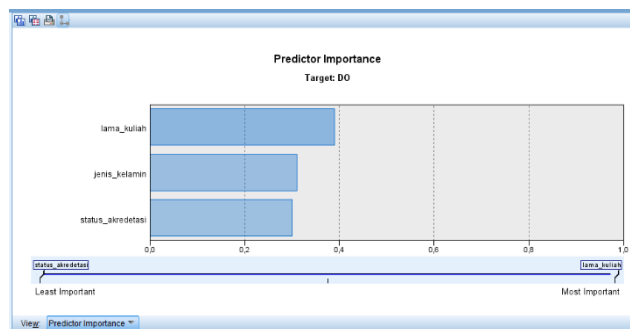
Gambar 3. Proses Modeling

Dari beberapa field yang digunakan pada proses pemodelan terdapat 3 *field* yang menjadi *predictor importance* / faktor yang paling mempengaruhi terhadap hasil prediksi, yaitu lama kuliah dengan nilai *predictor importance* sebesar 0,39, kemudian *predictor importance* selanjutnya adalah jenis_kelamin dengan nilai 0,31 dan *predictor importance* yang terakhir adalah status akreditasi. Setelah dilakukan *modeling* maka terbentuk pohon klasifikasi prediksi mahasiswa DO dengan algoritma *chaid*, terdiri dari 4 tahapan klasifikasi yaitu :

1. Pada klasifikasi pertama adalah penentuan DO atau tidak nya mahasiswa dengan nilai persentase sebesar 99,4 % untuk status mahasiswa tidak DO dan 0,55% untuk status mahasiswa DO.
2. Pada tahap klasifikasi yang kedua adalah *predictor importance* lama_kuliah yang diklasifikasikan menjadi dua kategori yaitu mahasiswa dengan lama kuliah kurang dari atau sama dengan 2 tahun dengan nilai persentase 98,98 % untuk status mahasiswa tidak DO 1,12 %, untuk status mahasiswa DO dan kategori mahasiswa dengan lama kuliah lebih dari 2 tahun.
3. Pada tahap klasifikasi yang ketiga adalah *predictor importance* status_akreditasi yang diklasifikasikan menjadi dua kategori yaitu blank atau data yang tidak terisi atau kosong pada field status_akreditasi dengan nilai persentase 98,08 % untuk status mahasiswa tidak DO dan 1,92 % untuk status mahasiswa DO dan kategori yang kedua adalah field yang sudah terisi akreditasi B, C dan Null.
4. Pada tahap klasifikasi yang keempat adalah *predictor importance* jenis_kelamin yang diklasifikasikan menjadi dua kategori yaitu mahasiswa berjenis kelamin pria berstatus DO dengan nilai persentase 98,72 % dan mahasiswa berjenis kelamin wanita berstatus DO dengan nilai persentase 1,28 % .

Dari hasil summary model tersebut terdiri dari *field* input sebagai *predictor importance* yang terdiri: (1) status_akreditasi; (2) jenis_kelamin; (3) lama_kuliah dan *field* target dari *summary model* adalah *field* DO.

Gambar 4 merupakan *summary model* pada spss modeler setelah dilakukan *modeling* dengan algoritma *Chaid* :



Gambar 4. Hasil *Summary Model*

Dari data modeling menghasilkan hasil klasifikasi dari mahasiswa program studi Manajemen Informatika, Sistem Informasi dan Teknik Informatika dengan jumlah prediksi mahasiswa DO 1.152 dan mahasiswa yang tidak DO 4.117. Gambar 5 adalah output data yang di peroleh :

1 tahun	2 tahun	hrs tahun	target	nim	jurusan	kelas	tgl lulus	lok	semester	nama	jenis_kelamin	kelstatus	nik_penanggu	Partition	SC-target	SCC-target
0	0	2010	0	06110023	SI-S1	KARYAWAN	0,0	0	0	Michael G Pria	Belum Me Sendiri	2_Testing	0	0,866289		
1	0	2010, 2008	1	06110024	SI-S1	KRYWN FT/2008	0,0	0	0	Agus Pram Pria	Belum Me Sendiri	1_Training	0	0,866289		
1	0	2010, 2010	1	06110025	SI-S1	MALAKA F/2010	3,46	5	5	Loly Rahm Wanita	Belum Me Crang tua	2_Testing	1	0,833333		
0	0	2010	0	06110026	SI-S1	PAGI FTW	0,0	0	0	Gangga Ds Pria	Belum Me Crang tua	1_Training	0	0,866289		
1	0	2010, 2009	1	06110027	SI-S1	FTW C 2009	3,1	5	5	Mohamad Pria	Belum Me Sendiri	2_Testing	1	0,915254		
1	0	2010, 2010	1	06110028	SI-S1	MALAKA F/2010	3,1	5	5	Husnul Ha Wanita	Belum Me Crang tua	1_Training	1	0,833333		
1	0	2010, 2010	1	06110029	SI-S1	FTW C 2010	3,29	7	7	Imam Suci Pria	Sudah Me Sendiri	1_Training	1	0,929663		
0	0	2010	0	06110030	SI-S1	MALAKA A	0,0	0	0	Najamud: Pria	Belum Me Crang tua	2_Testing	0	0,866289		
0	0	2010, 2011	0	06110031	SI-S1	FTW C 2011	3,1	10	10	Rizal Aprn: Pria	Belum Me Crang tua	1_Training	0	0,866289		
1	0	2010, 2010	1	06110032	SI-S1	FTW C 2010	3,0	7	7	Jedi Pekik Pria	Belum Me Crang tua	1_Training	1	0,952963		
1	0	2010, 2010	1	06110033	SI-S1	MALAKA F/2010	2,9	5	5	Masni Jun Wanita	Belum Me Crang tua	1_Training	1	0,833333		
1	0	2010, 2010	1	06110034	SI-S1	KARYAWA/2010	3,01	5	5	Fajar Kurn: Pria	Belum Me Sendiri	1_Training	1	0,915254		
1	0	2010, 2009	1	06110035	SI-S1	FTW C 2009	3,0	5	5	Ado Sutris: Pria	Belum Me Sendiri	1_Training	0	0,866289		
1	0	2010, 2010	1	06110036	SI-S1	MALAKA C/2010	3,04	5	5	Nila Maris Wanita	Belum Me Crang tua	2_Testing	1	0,825		
0	0	2010	0	06110037	SI-S1	MALAKA C	0,0	0	0	Wawan Ju: Pria	Sudah Me Sendiri	2_Testing	0	0,866289		
1	0	2010, 2010	1	06110038	SI-S1	MALAKA F/2010	3,1	5	5	Asep Hida: Pria	Belum Me Sendiri	2_Testing	0	0,926829		
1	0	2010, 2010	1	06110039	SI-S1	MALAKA F/2010	3,3	5	5	Fuji Farha Wanita	Belum Me Crang tua	1_Training	1	0,583333		
0	0	2010	0	06110040	SI-S1	MALAKA A	0,0	0	0	Ferhatsya: Pria	Belum Me Crang tua	1_Training	0	0,866289		
1	0	2010, 2008	1	06110041	SI-S1	PAGI 2008	3,03	4	4	Novie Put: Pria	Belum Me Crang tua	1_Training	0	0,751657		
1	0	2010, 2010	1	06110042	SI-S1	MALAKA F/2010	2,9	5	5	Suryani Wanita	Belum Me Crang tua	1_Training	1	0,833333		
0	0	2010	0	06110043	SI-S1	FTW A	0,0	0	0	Denyys Si: Pria	Belum Me Crang tua	1_Training	0	0,866289		
0	0	2010, 2011	0	06110044	SI-S1	FTW C 2011	3,05	5	5	Dwi Amie: Wanita	Belum Me Sendiri	2_Testing	0	0,926829		
0	0	2010	0	06110045	SI-S1	MALAKA C	0,0	0	0	Endar Pam: Pria	Belum Me Sendiri	2_Testing	0	0,866289		
0	0	2010	0	06110046	SI-S1	PAGI FTW	0,0	0	0	Ahmad Az: Pria	Belum Me Crang tua	1_Training	0	0,866289		

Gambar 5. Output data yang di olah IBM SPSS menggunakan algoritma *chaid*

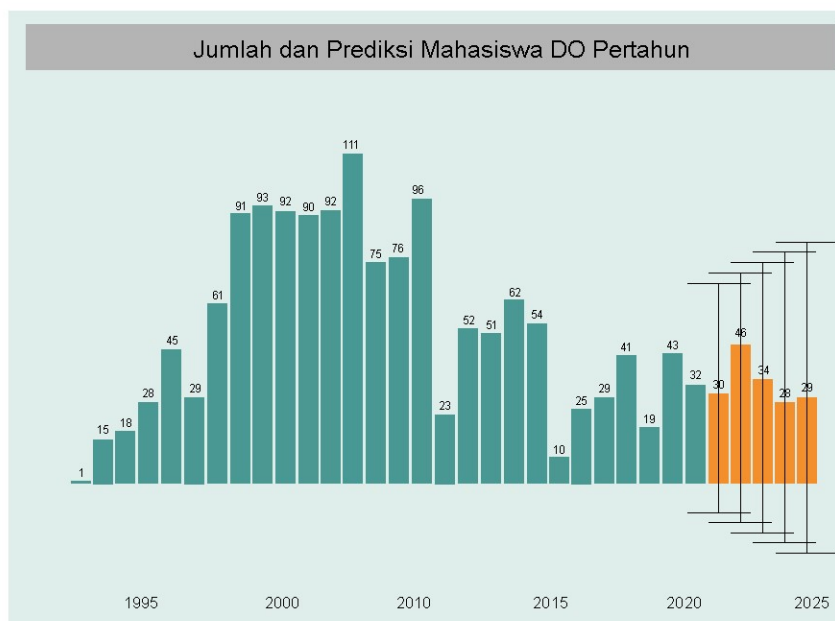
Agar data prediksi mahasiswa DO yang telah dimodeling mudah untuk dipahami oleh pimpinan dan ketua jurusan perlu dilakukan visualisasi data ke dalam sebuah diagram atau tampilan yang mudah dipahami oleh orang awam, dengan memanfaatkan *tableau application* data akademik divisualisasikan berdasarkan kebutuhan informasi yang akan ditampilkan.

Pada gambar 6 memperlihatkan kenaikan dan penurunan jumlah mahasiswa IBT Swadharma jurusan studi Sistem Informasi, Teknik Informatika dan Manajemen Informatika secara keseluruhan dalam rentan waktu 1992-2019 :



Gambar 6. Total Mahasiswa

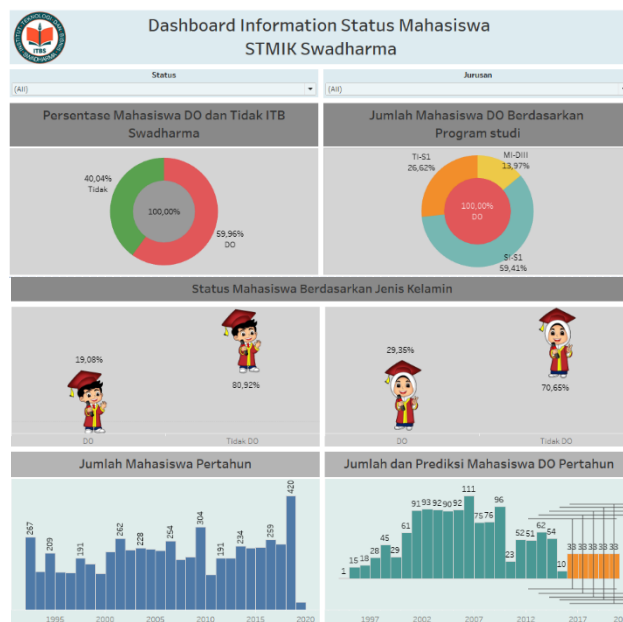
Berikut visualisasi jumlah mahasiswa DO dan prediksi mahasiswa DO selama 5 tahun yaitu antara tahun 2020 sampai tahun 2025, dengan menggunakan *data training* jumlah mahasiswa DO dari tahun 1992 sampai 2019, dengan melihat laju pertumbuhan jumlah mahasiswa pertahun sehingga dapat diprediksi jumlah mahasiswa DO seperti terlihat pada gambar 7 :



Gambar 7. Prediksi Mahasiswa DO

Salah satu teknologi yang dapat memberikan solusi terhadap kebutuhan penyajian data dan informasi kedalam bentuk visual yang menarik dan mudah untuk dimengerti adalah dashboard. Tampilan yang memungkinkan para pemegang keputusan mendapatkan ringkasan informasi penting secara cepat dan mudah

sebagai landasan untuk menetapkan keputusan penting bagi perusahaan. Berikut *dashboard* prediksi mahasiswa DO program studi Manajemen Informatika, Sistem Informasi dan teknik Informatika di ITB Swadharna, *dashboard* berupa gabungan dari beberapa visualisasi yang telah dibuat. Gambar 8 adalah tampilan *dashboard* yang dibuat :



Gambar 8. Dashboard Prediksi Mahasiswa Putus Kuliah (DO)

IV. Penutup

Dari hasil pembahasan didapati sejumlah fakta tentang prediksi mahasiswa DO sebagai berikut :

1. Data diolah adalah data yang didapat sejak 1992 hingga 2019 dengan data perolehan 5.840. Data yang sudah diolah menghasilkan jumlah signifikan faktor terhadap prediksi mahasiswa DO dan mahasiswa yang tidak DO. Total prediksi pada mahasiswa DO dengan presentase 0.559 % sejumlah 19. Sedangkan prediksi mahasiswa yang tidak DO dengan presentase 99.441 % sejumlah 3378 dari total keseluruhan data.
2. Signifikan faktor yang mempengaruhi disekitarnya antara lain : Faktor lama kuliah menjadi faktor utama yang berpengaruh terhadap prediksi mahasiswa DO dengan persentase 1.019 % sejumlah 19, dan P_value 0.000 dengan nilai chi-square 15.714 . Kemudian faktor status akreditasi dengan presentase 1.915 % sejumlah 19, dengan jumlah p_value 0.000 dan nilai chi-square 16.874. faktor yang berpengaruh lainnya adalah faktor jenis kelamin dengan presentase 1.277 % sejumlah 9, dengan jumlah p-value 0.021 dan nilai chi-square 5.292.

Jadi dapat disimpulkan bahwa faktor utama mahasiswa putus kuliah pada ITB Swadharna adalah faktor lama kuliah, perlu dilakukan penelitian lebih lanjut tentang faktor – faktor penyebab lama kuliah yang terjadi. Penelitian dilakukan pada ITB Swadharna yang berlokasi di DKI Jakarta, perlu dilakukan penelitian pada universitas lain dengan demografi yang berbeda - beda untuk menemukan faktor lain yang bisa mempengaruhi angka mahasiswa DO.

Daftar Pustaka

- [1] H. Cahyono, “Peran Mahasiswa Di Masyarakat,” *J. Pengabd. Masy. setia budhi*, vol. 1, no. 1, pp. 32–43, 2019. Available at: <https://stkipsetiabudhi.e-journal.id/DeBode>
- [2] D. F. D. Putra, “Evaluasi Dan Pengembangan Knowledge Management System Menggunakan Post-Study System Usability Questionnaire(PSSUQ),” *J. Rekayasa Inf. Swadharna(JRIS)*, vol. 01, no. 01, pp. 9–13, 2021. Available at: <http://ejurnal.swadharna.ac.id/index.php/jris/article/view/58>
- [3] PDDIKTI Kemenristekdikti, *Statistik Pendidikan Tinggi Tahun 2017*. Jakarta: Pudatin Iptek Dikti Setjen Kemenristekdikti, 2017.

- [4] Shaw, E. J., & Mattern, K. D. (2013). Examining Student Under- and Overperformance in College to Identify Risk of Attrition. *Educational Assessment*, 18(4), 251–268.
- [5] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, Third. Waltham, MA: Morgan Kaufmann, 2012.
- [6] B. G. Sudarsono and A. U. Bani, “Prediksi Mahasiswa Berpotensi Berhenti Kuliah Secara Sepihak Menggunakan Data Mining Algoritma C.45,” *J. Sains Komput. Inform.*, vol. 4, no. 2, pp. 359–367, 2020.
- [7] A. P. U. Sembiring and M. Ginting, “Analisa Faktor-Faktor Yang Mempengaruhi Pengunduran Diri Mahasiswa dengan Aplikasi Data Mining Add-Ins Studi Kasus Pada STMIK Mikroskil,” *J. Sifo Mikroskil*, vol. 14, no. 2, 2013.
- [8] E. A. Hasibuan and A. N. Harahap, “Aplikasi Metode Chaid dalam Menganalisis Kecendrungan Penelitian Skripsi Mahasiswa pada Program Studi Pendidikan Matematika,” *Edumatika J. Ris. Pendidik. Mat.*, vol. 1, no. 2, pp. 63–72, 2018.
- [9] Y. S. Kunto and S. N. Hasana, “ANALISIS CHAID SEBAGAI ALAT BANTU STATISTIKA UNTUK SEGMENTASI PASAR (Studi Kasus pada Koperasi Syari’ah Al-Hidayah),” *J. Manaj. Pemasar.*, vol. 1, no. 2, pp. 88–97, 2006. Available at: <http://puslit2.petra.ac.id/ejournal/index.php/mar/article/view/16538>.
- [10] A. Dharmalau, L. Nurlaela, and M. Hartono, “Perancangan sistem informasi paguyuban alumni STMIK Swadharma berbasis web,” *Syntax Idea*, vol. 2, no. 4, pp. 12–25, 2020.
- [11] Annisa Paramitha Fadillah, " Penerapan Metode CRISP-DM untuk Prediksi Kelulusan Studi Mahasiswa Menempuh Mata Kuliah (Studi Kasus Universitas XYZ)," *Jurnal Teknik Informatika dan Sistem Informasi* e-ISSN : 2443-2229 Volume 1 Nomor 3 Desember 2015
- [12] Y. Suhandi, I. Kurniati, and S. Norma, “Penerapan Metode Crisp-DM dengan Algoritma K-Means Clustering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik,” *J. Teknol. Inform. dan Komput. MH Thamrin*, vol. 6, no. 2, pp. 12–20, 2020. doi: <https://doi.org/10.37012/jtik.v6i2.299>.
- [13] Mariscal, Gonzalo, Oscar Marban, and Covadonga Fernandez. 2010. “A Survey of Data Mining and Knowledge Discovery Process Models and Methodologies.” *The Knowledge Engineering Review*, Cambridge University Press 25:2(2010): 137–66.