

Karakteristik Kelulusan Tepat Waktu Mahasiswa FMIPA UNP dengan Menggunakan Metode *Classification and Regression Trees*

Lilis Sulistiawati^{#1}, Nonong Amalita^{*2}

Mathematics Department Universitas Negeri Padang, Indonesia

Jl. Prof. Hamka Air Tawar, Padang, Indonesia

¹sulistiawatililis757@gmail.com

²nongmat@fmipa.unp.ac.id

Abstract — This reaserch discussed about classification of timely graduation of college students. There are still a lot of students of Mathematics and Natural Science Faculty of Universitas Negeri Padang who was overdue in Finishing their studies. This can affect the decline in study program accreditation or even reduce the university's image. Therefore, this reaserch aims to determine the characteristic of students who graduate on time and who do not graduate on time based on variables of gender, school origin, scholarship, college entrance pathway, Grade Point Average (GPA), father' occupation and mother's occupation by using the Classification and regression Trees (CART) method. The population in this reaserch is the entire student of Mathematics and Natural Science Faculty of Universitas Negeri Padang who graduated in 2018. The characteristics of college students who graduate on time is students who have a 3,51 – 4,00 GPA with a percentage of 75,2% and student who have a 2,00 – 3,50 GPA, received scholarship and came through the SBMPTN with a percentage 83,3%.

Keywords — Characteristic of Graduation, Classification and Regression Trees (CART), Optimum Tree.

Abstrak — Penelitian ini membahas tentang klasifikasi karakteristik kelulusan tepat waktu mahasiswa. Masih banyak mahasiswa FMIPA UNP yang mengalami keterlambatan dalam menyelesaikan studinya. Hal ini dapat berpengaruh terhadap menurunnya akreditasi program studi atau bahkan menurunkan citra universitas. Oleh karena itu, penelitian ini bertujuan untuk mengetahui ciri kelompok mahasiswa yang lulus tepat waktu dan yang tidak lulus tepat waktu berdasarkan variabel jenis kelamin, daerah asal sekolah, beasiswa, jalur masuk, Indeks Prestasi Kumulatif (IPK), pekerjaan ayah dan pekerjaan ibu dengan menggunakan metode *Classification and Regression Trees* (CART). Populasi pada penelitian ini adalah seluruh wisudawan FMIPA UNP yang lulus pada tahun 2018. Karakteristik mahasiswa yang lulus tepat waktu adalah mahasiswa yang memiliki IPK 3,51 – 4,00 dengan persentase 75,2% dan mahasiswa yang memiliki karakteristik IPK 2,00 – 3,50, menerima beasiswa dan masuk melalui jalur SBMPTN dengan persentase 83,3%.

Kata kunci — Karakteristik Kelulusan, *Classification and Regression Trees* (CART), Pohon Optimum.

PENDAHULUAN

Perguruan tinggi adalah jenjang pendidikan tertinggi di Indonesia yang memiliki tujuan untuk menghasilkan profesional-profesional yang ahli dibidangnya guna meningkatkan daya saing bangsa. Untuk menyelesaikan pendidikan di perguruan tinggi dibutuhkan waktu perkuliahan selama masa yang sudah ditentukan.

Salah satu indikator keberhasilan mahasiswa dalam menempuh pendidikan di perguruan tinggi adalah lulus tepat waktu. Mahasiswa dikatakan lulus tepat waktu apabila dapat menyelesaikan studinya selama tiga tahun untuk program Diploma III (D3) dan 3,5 sampai empat tahun untuk program Strata I (S1). Namun sayangnya masih banyak mahasiswa yang menyelesaikan studinya

lewat dari masa tersebut, sehingga dikatakan tidak lulus tepat waktu.

Permasalahan ini juga dihadapi oleh Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Negeri Padang (UNP). Sampai saat ini mahasiswa FMIPA UNP masih belum sepenuhnya mampu menyelesaikan studi sesuai jadwal yang telah ditentukan. Berdasarkan data wisudawan FMIPA UNP yang lulus pada tahun 2016 – 2018, mahasiswa yang tidak lulus tepat waktu ada sebanyak 43% dari total keseluruhan. Hal ini belum sesuai dengan tujuan fakultas yang ingin menghasilkan lulusan yang berkualitas dari waktu ke waktu dan dapat bersaing ditingkat nasional.

Setiap mahasiswa tentu berharap agar studinya berjalan lancar, lulus tepat waktu, dan memiliki prestasi

akademik yang tinggi. Namun, pada kenyataannya banyak ditemui hambatan atau bahkan keterlambatan yang dapat berpengaruh terhadap beban biaya studi hingga pada citra universitas. Semakin banyak mahasiswa yang mengalami keterlambatan, maka akan semakin berpengaruh terhadap penurunan akreditasi program studi di universitas tersebut. Apabila hal ini terjadi pada mahasiswa yang memiliki motivasi belajar yang rendah, dikhawatirkan akan menghambat masa studinya dan terjadi *Drop Out* (DO).

Banyak faktor yang mempengaruhi hasil belajar mahasiswa, salah satunya adalah faktor prestasi belajar. Faktor prestasi belajar mahasiswa terdiri dari faktor internal dan faktor eksternal [1]. Faktor yang berasal dari dalam diri suatu individu seperti faktor fisik dan faktor psikis merupakan faktor internal, sedangkan faktor yang berasal dari luar individu seperti faktor keluarga, sekolah, maupun faktor lingkungan sosial lainnya merupakan faktor eksternal. Faktor lain yang mempengaruhi hasil belajar mahasiswa ada dua yaitu faktor utama seperti kemampuan, minat, motivasi, dan IPK atau hasil belajar sebelumnya. Lalu faktor kedua yaitu karakter pengajar dan fasilitas belajar [2]. Selain itu, mahasiswa perempuan juga cenderung memiliki prestasi akademik yang lebih baik dibandingkan dengan laki-laki karena mereka lebih tekun, teliti, dan mendengarkan dengan baik [3].

Ada beberapa hal yang dapat memotivasi mahasiswa dalam menyelesaikan studinya tepat waktu, salah satunya adalah beasiswa. Dengan pemberian beasiswa maka ada keringanan biaya studi yang didapat oleh mahasiswa, hal ini dapat memotivasi mereka agar selalu berprestasi dan menyelesaikan studinya tepat waktu. Bahkan salah satu jenis beasiswa yaitu bidikmisi, hanya diberikan pada mahasiswa kurang mampu secara finansial sampai pada batas waktu yang sudah dijadwalkan. Karakteristik penerima beasiswa bidikmisi kebanyakan yang memiliki ayah seorang petani, tukang ojek atau bahkan tidak punya pekerjaan dengan luas rumah tidak lebih dari 50 m²[4]. Hal ini juga memotivasi mereka untuk dapat meringankan beban orangtua dengan lulus secepatnya.

Untuk mengantisipasi masalah ini, perlu diketahui karakteristik mahasiswa yang dapat mencirikan/menjelaskan perbedaan mahasiswa yang lulus tepat waktu. Berdasarkan faktor-faktor diatas, maka variabel yang digunakan dalam penelitian ini adalah jenis kelamin, daerah asal sekolah, jalur masuk, beasiswa, IPK, dan pekerjaan orangtua. Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengklasifikasikan karakteristik-karakteristik yang paling berpengaruh terhadap waktu kelulusan. Serta mengelompokkan mahasiswa FMIPA UNP, dimana nantinya akan bermanfaat dalam meminimalisir jumlah keterlambatan mahasiswa dalam menyelesaikan studi.

Untuk melakukan klasifikasi tersebut, metode yang digunakan adalah metode *Classification and Regression Trees* (CART). Metode CART adalah salah satu metode statistik nonparametrik khusus untuk topik analisis klasifikasi yang memiliki dua tipe analisis, yaitu analisis pohon klasifikasi untuk variabel respon bertipe kategorik,

dan analisis pohon regresi untuk variabel respon bertipe kontinu [5]. Metode ini memiliki tujuan utama untuk memperoleh suatu kelompok data yang akurat sebagai penciri suatu kelompok pengklasifikasian.

METODE

Penelitian ini merupakan penelitian terapan. Jenis data yang digunakan adalah data sekunder, yaitu data wisudawan FMIPA UNP yang lulus pada tahun 2018 dengan jumlah data sebanyak 625. Data diperoleh dari UPT Puskom UNP.

Variabel-variabel pada penelitian ini terdiri dari variabel terikat dan variabel bebas. Variabel terikat (Y) berskala nominal yang digunakan adalah status kelulusan yang dikelompokkan menjadi dua, yaitu:

- Y = 0, untuk kategori tidak tepat waktu.
- Y = 1, untuk kategori tepat waktu.

Variable bebas (X) yang digunakan adalah:

1. Jenis kelamin (X_1), berskala nominal
 - $X_1 = 0$, untuk kategori laki-laki.
 - $X_1 = 1$, untuk kategori perempuan.
2. Daerah asal sekolah (X_2), berskala nominal
 - $X_2 = 0$, untuk kategori desa.
 - $X_2 = 1$, untuk kategori kota.
3. Jalur masuk (X_3), berskala nominal
 - $X_3 = 0$, untuk kategori mandiri.
 - $X_3 = 1$, untuk kategori jalur prestasi.
 - $X_3 = 2$, untuk kategori SBMPTN.
 - $X_3 = 3$, untuk kategori SNMPTN.
4. Beasiswa (X_4), berskala nominal
 - $X_4 = 0$, untuk kategori tidak menerima.
 - $X_4 = 1$, untuk kategori menerima.
5. IPK (X_5), berskala ordinal
 - $X_5 = 0$, untuk kategori 2,00 – 2,75.
 - $X_5 = 1$, untuk kategori 2,76 – 3,50.
 - $X_5 = 2$, untuk kategori 3,51 – 4,00.
6. Pekerjaan ayah (X_6), berskala nominal
 - $X_6 = 0$, untuk kategori lain-lain.
 - $X_6 = 1$, untuk kategori pensiunan.
 - $X_6 = 2$, untuk kategori petani/nelayan.
 - $X_6 = 3$, untuk kategori wiraswasta.
 - $X_6 = 4$, untuk kategori karyawan swasta.
 - $X_6 = 5$, untuk kategori PNS.
7. Pekerjaan ibu (X_7), berskala nominal
 - $X_7 = 0$, untuk kategori lain-lain.
 - $X_7 = 1$, untuk kategori pensiunan.
 - $X_7 = 2$, untuk kategori petani/nelayan.
 - $X_7 = 3$, untuk kategori wiraswasta.
 - $X_7 = 4$, untuk kategori karyawan swasta.
 - $X_7 = 5$, untuk kategori PNS.

Analisis data dilakukan dengan menggunakan bantuan *Software Salford Predictive Modeller* (SPM) 7.0. langkah-langkah yang dilakukan adalah sebagai berikut:

1. Pemilihan variabel terbaik yang digunakan untuk memecah sebuah simpul dengan kriteria nilai *improvement* atau nilai perubahan *impurity* tertinggi.

2. Memberi label pada kelas yang memiliki proporsi tertinggi.
3. Penghentian pembentukan pohon klasifikasi karena pada ujung pohon terdapat simpul yang memiliki anggota pada kelas yang sama (homogen).
4. Setelah pohon maksimum terbentuk, maka akan dilakukan pemangkasan pada simpul induk dan kedua simpul anaknya yang memenuhi persamaan $R(t) = R(t_R) + R(t_L)$ hingga menghasilkan pohon yang jauh lebih sederhana.
5. Pemilihan atau penentuan pohon klasifikasi optimal berdasarkan *subtree* yang memiliki nilai *cross validation relative cost* terkecil.
6. Menguji keakuratan pohon klasifikasi optimal.
7. Menginterpretasikan hasil dari pohon klasifikasi metode CART.

HASIL DAN PEMBAHASAN

A. Hasil Penelitian

Berdasarkan data wisudawan yang lulus tahun 2018 sebanyak 625 orang yang diperoleh dari UPT Puskom UNP, diketahui bahwa wisudawan yang lulus tepat waktu hanya sekitar 45,76% dari total keseluruhan. Diketahui juga bahwa yang mendominasi kelulusan adalah wisudawan yang memiliki karakteristik jenis kelamin perempuan (85,92%), daerah asal sekolah dari desa (52,64%), jalur masuk SNMPTN (47,84%), tidak menerima beasiswa (66,72%), IPK antara 2,76 – 3,50 (80,16%), pekerjaan ayah petani/nelayan (27,36%), dan pekerjaan ibu lain-lain (61,28%).

1. Proses pemecahan simpul

Pemecahan akan dilakukan pada simpul induk yang dipilih berdasarkan pada kriteria *goodness of split* (pemecah terbaik). Kriteria ini dihitung menggunakan fungsi nilai *impurity* tertinggi dari setiap variabel yang ada. Variabel IPK terpilih sebagai simpul induk yang akan dipecah karena memiliki *improvement* tertinggi dengan perhitungan :

$$\begin{aligned} \Delta i(s, t) &= \left(1 - \left[\left(\frac{286}{625} \right)^2 + \left(\frac{339}{625} \right)^2 \right] \right) \\ &\quad - \left[\left(\frac{508}{625} \right) \left(1 - \left[\left(\frac{198}{508} \right)^2 + \left(\frac{310}{508} \right)^2 \right] \right) \right] \\ &\quad - \left[\left(\frac{117}{625} \right) \left(1 - \left[\left(\frac{88}{117} \right)^2 + \left(\frac{29}{117} \right)^2 \right] \right) \right] \\ &= 0,03905 \end{aligned}$$

Pemecahan terhadap simpul induk (*Node 1*) menghasilkan dua simpul anak (*Terminal node*) dimana *Terminal Node 1* terbentuk berdasarkan kriteria variabel IPK 2,00 – 2,75; 2,76 – 3,50.. Sedangkan *Terminal Node 2* terbentuk berdasarkan kriteria variabel IPK 3,51 – 4,00. Pemecahan ini dilakukan secara terus menerus hingga tidak memungkinkan lagi terjadinya pemecahan atau simpul akhir bersifat homogen.

2. Pelabelan kelas

Simpul-simpul yang terbentuk diberi label sesuai kelas dengan proporsi terbesar. Proses pelabelan pada simpul induk dapat dilihat sebagai berikut:

$$P(\text{tepat waktu} | t) = \frac{286}{625} = 0,4576$$

$$P(\text{tidak tepat waktu} | t) = \frac{339}{625} = 0,5424$$

Karena proporsi kelas tidak tepat waktu lebih besar, maka simpul induk diberi label kelas tidak tepat waktu. Proses ini dilakukan pada semua simpul terutama simpul terminal, karena simpul terminal sangat penting untuk memprediksi suatu objek yang ada pada kelas tertentu.

3. Proses penghentian pemecahan

Proses pemecahan simpul yang dilakukan berulang-ulang menghasilkan 67 simpul nonterminal, dan 68 simpul terminal. Penghentian pemecahan terjadi saat diujung pohon klasifikasi terdapat simpul dengan anggota kelas yang sama (homogen).

4. Proses pemangkasan pohon

Proses pemangkasan pohon dilakukan karena pohon maksimum yang dihasilkan sangat besar. Setiap pemangkasan suatu simpul akan menghasilkan pohon baru yang lebih sederhana (*Subtree*). Dari proses ini, diperoleh 24 *subtree* yang nantinya akan dipilih satu sebagai pohon klasifikasi optimal.

5. Pohon klasifikasi optimal

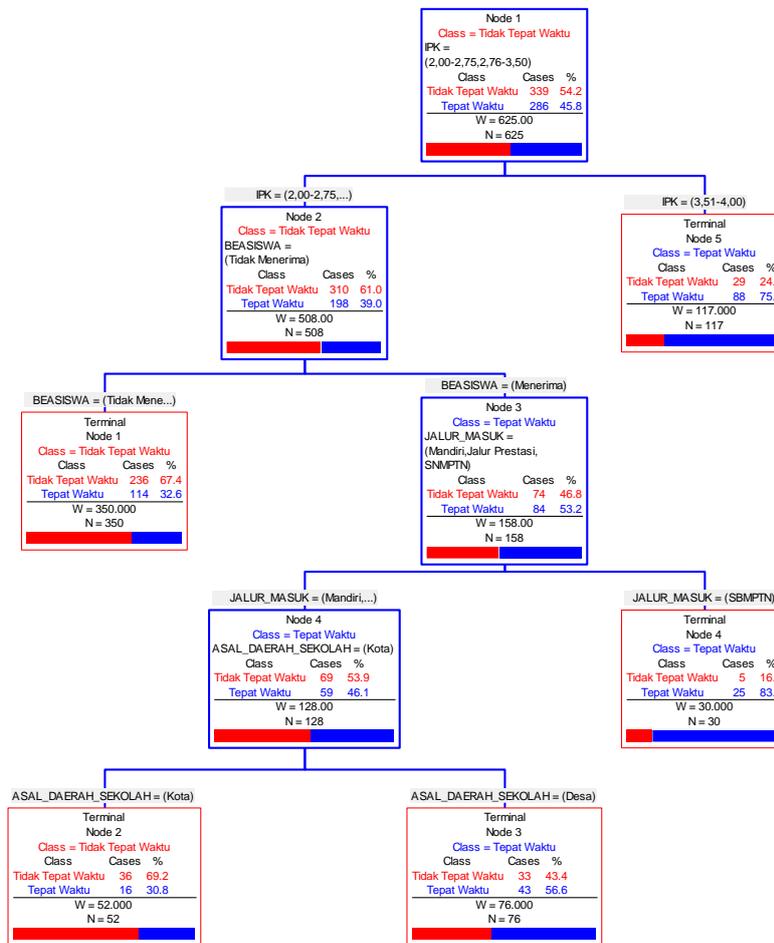
Setelah proses pemangkasan pohon klasifikasi yang menghasilkan 24 *subtree*, maka akan dipilih satu *subtree* dengan nilai penduga yang terkecil.

TABEL 1
Tree Sequence pada SPM 7.0

<i>Tree</i>	<i>Terminal Nodes</i>	<i>CrossValid Rel Error</i>	<i>Resub Rel Error</i>
1	68	0,82830	0,50893
2	66	0,83562	0,51003
3	65	0,83803	0,51079
4	62	0,83398	0,51319
5	58	0,83344	0,51669
6	56	0,82808	0,51866
7	52	0,82164	0,52303
8	48	0,81923	0,52805
9	42	0,82994	0,53614
10	37	0,81355	0,54313
11	36	0,81464	0,54477
12	32	0,81060	0,55243
13	28	0,81115	0,56128
14	25	0,79290	0,56882
15	24	0,76493	0,57177
16	23	0,76012	0,57493
17	19	0,74963	0,58804
18	18	0,73914	0,59154
19	12	0,72276	0,61339
20	6	0,70474	0,64334
21	5	0,67447*	0,65219
22	3	0,70244	0,70244
23	2	0,77785	0,77785
24	1	1,00000	1,00000

Berdasarkan Tabel 1, terlihat bahwa *subtree* 21 memiliki nilai penduga terkecil dari *subtree* lainnya,

yaitu sebesar 0,67447. Subtree ini menghasilkan Sembilan simpul dengan lima *Terminal Nodes*.



Gambar 1. Pohon Klasifikasi Optimal pada SPM 7.0

6. Keakuratan Klasifikasi

Pohon klasifikasi optimal yang telah diperoleh akan diuji keakuratannya dalam mengklasifikasikan kelulusan tepat waktu mahasiswa FMIPA UNP tahun 2018.

TABEL 2
Confusion Matrix pada SPM 7.0

Actual Class	Total Class	Percent Correct	Predicted Classes	
			Tidak Tepat Waktu N = 401	Tepat Waktu N = 224
Tidak Tepat Waktu	339	79,06%	268	71
Tepat Waktu	286	53,50%	133	153
Total	625			
Average		66,28%		
Overall % Correct		67,36%		
Specificity		79,06%		
Sensitivity		53,50%		

Berdasarkan Tabel 2, nilai keakuratan pohon klasifikasi optimal adalah sebesar 67,36%.

B. Pembahasan

Dari hasil analisis data menggunakan metode CART, diperoleh empat variabel yang berpengaruh secara signifikan terhadap waktu kelulusan mahasiswa FMIPA UNP yang lulus pada tahun 2018. Variabel tersebut adalah IPK, beasiswa, jalur masuk, dan daerah asal sekolah.

Hal ini juga dapat dilihat dalam bentuk pohon klasifikasi optimal yang memiliki 4 kedalaman dimana variabel IPK membagi waktu lulus pada kedalaman ke-1, variabel beasiswa membagi waktu lulus pada kedalaman ke-2, variabel jalur masuk membagi waktu lulus pada kedalaman ke-3, dan variabel daerah asal sekolah membagi waktu lulus pada kedalaman ke-4.

Karena penghentian pemecahan terjadi pada simpul terminal 1, 2, 3, 4, dan 5, maka kelas kelulusan tepat waktu mahasiswa dapat diklasifikasikan menjadi lima kelompok pengklasifikasian.

TABEL 3
Hasil Klasifikasi Pohon Optimal dengan Metode CART

	Status Kelulusan	Terminal Node	Persentase	Karakteristik
Klasifikasi Kelulusan	Tidak tepat Waktu	1	67,4%	Mahasiswa yang memiliki IPK 2,00 – 3,50 dan tidak menerima beasiswa
		2	69,2%	Mahasiswa yang memiliki IPK 2,00 – 3,50, menerima beasiswa, masuk melalui jalur mandiri, prestasi, atau SNMPTN, serta daerah asal sekolah dari kota
	Tepat Waktu	3	56,6%	Mahasiswa yang memiliki IPK 2,00 – 3,50, menerima beasiswa, masuk melalui jalur mandiri, prestasi, atau SNMPTN, serta daerah asal sekolah dari desa
		4	83,3%	Mahasiswa yang memiliki IPK 2,00 – 3,50, menerima beasiswa, dan masuk melalui jalur SBMPTN
		5	75,2%	Mahasiswa yang memiliki IPK 3,51 – 4,00

Berdasarkan Tabel 3, diketahui klasifikasi kelulusan tepat waktu tertinggi adalah kelompok pada *node* 4, yaitu mahasiswa yang memiliki karakteristik IPK 2,00 - 3,50, menerima beasiswa, dan jalur masuk melalui jalur SBMPTN dengan persentase 83,3%. Namun kelompok pada *node* 5 juga penting, karena hanya dengan karakteristik IPK 3,51 – 4,00 kelompok ini memiliki persentase kelulusan tepat waktu sebesar 75,2%. Oleh karena itu, dapat disimpulkan ciri mahasiswa yang lulus tepat waktu adalah mahasiswa yang memiliki IPK 3,51 – 4,00. Namun jika mahasiswa tersebut tidak memiliki IPK 3,51 – 4,00 maka diperlukan karakteristik menerima beasiswa dan dan masuk melalui jalur SBMPTN.

SIMPULAN

Berdasarkan pohon klasifikasi optimal tentang karakteristik kelulusan tepat waktu mahasiswa FMIPA UNP dengan menggunakan metode CART, pengklasifikasian terbagi menjadi lima kelompok karakteristik dengan variabel yang paling berpengaruh terhadap waktu kelulusan mahasiswa adalah mahasiswa

yang memiliki variabel IPK, beasiswa, jalur masuk, dan daerah asal sekolah.

Hasil dari analisis menggunakan metode CART diperoleh bahwa karakteristik mahasiswa yang lulus tepat waktu adalah mahasiswa yang memiliki IPK 3,51 – 4,00 dengan persentase 75,2% dan mahasiswa yang memiliki IPK 2,00 – 3,50, menerima beasiswa dan masuk melalui jalur SBMPTN dengan persentase 83,3%.

REFERENSI

- [1] Slameto. 2003. *Belajar dan Faktor-Faktor yang Mempengaruhinya*. Jakarta: Rineka Cipta.
- [2] Budiningsih, Asri. 2005. *Belajar dan Pembelajaran*. Jakarta: Rineka Cipta.
- [3] Nuryoto, Sartini. 1998. Perbedaan Prestasi Akademik Antara Laki-laki dan Perempuan Studi di Wilayah Yogyakarta. *Jurnal Psikologi*, ISSN: 0215-8884 Nomor 2. Hlm:16-24.
- [4] Amalita, N., Kurniawati, Y., & Fitria, D. Classification Bidikmisi Scholarship Awardee in FMIPA UNP Using Biplot Analysis Mathematics and Mathematics Education (ICM2E), 242
- [5] Breiman et al. 1993. *Classification and Regression Trees*. New York: Chapman and Hall.