

Analysis of Academic Characteristics of USK FMIPA Undergraduate Students Using the Chaid Method

Fitriana AR^{1*}, Latifah Rahayu Siregar¹, Arif Yunandar¹

¹ Statistics Department, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh, 23111, Indonesia

*Corresponding email: fithriana.ar@fmipa.usk.ac.id

ARTICLE INFO	ABSTRACT
Article History Received : 01-08-2023 Revised : 27-10-2023 Accepted : 29-10-2023 Keyword Study period; Chi-square test; tree diagram; graduate students; classification.	The purpose of this research is to find out the most significant factors to the student's study period, to divide the data into smaller groups based on the relationship between the response variables and the explanatory variables. The characteristics and educational background are gender, school origin, major, supervisor's level of education, Grade Point Average (GPA), and trial grades. The method displays a graphical tree structure in the form of a diagram that explains the relationship between categorical response variables and categorical explanatory variables so that it is easy to interpret. The test used to produce a tree diagram is the Chi-square test and Bonferroni correction. In this study, data were used on graduates of the faculty of Mathematics and Natural Sciences USK class of 2017 who graduated from september 2020 to August 2022. The result of the study showed that the most significant factors for the period of study were GPA, school origin, supervisor education, level, type gender, and major. In terms of classification accuracy, the percentage accuracy of the model correctly classifying the study period is 71,8%.
DOI: ...	©2023 TJoMA. All rights reserved.

1. PENDAHULUAN

Mahasiswa adalah seseorang yang dianggap mampu memberikan perubahan kepada suatu negara/bangsa (*Agent of Change*) dan dianggap sebagai individu dengan tingkat intelektual yang baik. Usia saat menjadi mahasiswa adalah usia yang matang dalam menimba ilmu demi kemajuan bangsa serta generasi yang akan datang. Untuk menyelesaikan fase kematangan tersebut, ada batasan waktu yang disebut masa studi. Masa studi merupakan proses pembelajaran yang harus ditempuh oleh mahasiswa dengan rentang waktu yang telah ditentukan. Masa waktu studi merupakan batas waktu yang diperbolehkan oleh seorang mahasiswa dalam menyelesaikan perkuliahannya menurut kredit/beban SKS yang diambil mahasiswa tersebut. Masa waktu program studi Strata 1 (S1) yaitu tujuh tahun. Secara ideal mahasiswa S1 menyelesaikan program studinya yaitu sekitar empat sampai lima tahun.

Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Syiah Kuala (FMIPA USK) adalah satu dari banyaknya fakultas yang ada di Universitas Syiah Kuala. FMIPA USK memiliki beberapa program studi S1 yaitu prodi Matematika, Fisika, Kimia, Biologi, Statistika, Informatika, dan Farmasi. FMIPA USK berdiri sejak tahun 1989 dan telah meluluskan alumni yang berkompeten di bidangnya, beban studi program S1 FMIPA USK sekurang-kurangnya 144 sks dan maksimal 160 sks. Berkaitan dengan waktu studi mahasiswa, pada kenyataannya, berdasarkan data masa studi program S1 FMIPA USK, masih banyak mahasiswa yang tidak lulus tepat waktu. Faktor-faktor yang memengaruhi keterlambatan lulus kuliah diantaranya adalah faktor eksternal dan internal dari pribadi mahasiswa, salah satu faktor eksternal keterlambatan lulus kuliah ialah kerja dan salah jurusan, sedangkan faktor internal keterlambatan lulus kuliah ialah sikap mahasiswa yang

belum dewasa seperti tidak mengikuti perkuliahan dengan baik, tidak mengikuti peraturan yang di berikan oleh dosen (Tim Penyusun Kurikulum FMIPA, 2010).

Chi-Squared Automatic Interaction Detection (CHAID) digunakan untuk menampilkan struktur pohon yang terdiri dari diagram yang menjelaskan keterkaitan antara peubah respon kategorik dengan peubah penjelas kategorik. Metode CHAID biasanya diketahui sebagai metode pohon klasifikasi (*Classification Tree Method*). Metode tersebut membagi data kelompok menjadi lebih kecil (akar-akar pohon) dengan keterkaitan antara peubah respon dengan peubah penjelas. Metode CHAID secara harfiah bekerja menjelaskan hubungan antar peubah. Metode CHAID secara garis besar menjelaskan hubungan antara peubah respon dengan beberapa peubah penjelas selanjutnya mengklasifikasikan sampel berdasarkan keterkaitan tersebut (Agresti, 2002).

Hasil penelitian yang dilakukan oleh Fitri (2018) dengan sampel 88 orang menunjukkan bahwa mencapai target akademik yaitu 4,415 tahun. Analisis menggunakan metode CHAID menunjukkan bahwa ada faktor yang paling dominan yaitu waktu pengerjaan skripsi dan IPK yang mempengaruhi waktu kelulusan mahasiswa. Dalam hal ini akurasi yang dihasilkan sebesar 72,7%.

Aida (2015) menyatakan bahwa lulusan mahasiswa FMIPA Matematika USK (Januari 2007 – Agustus 2014) yang berjumlah 1034 orang menjadi tiga kategori berdasarkan masa studi. Setelah menganalisis dengan menggunakan metode CHAID, hasil penelitian menjelaskan bahwa karakteristik mahasiswa dengan masa studi cepat setara dengan masa studi sedang, ialah mahasiswa yang memperoleh IPK sangat memuaskan (*cumlaude*), nilai sidang A, dan jenis kelamin perempuan. Sedangkan karakteristik mahasiswa yang masa studi lama ialah mahasiswa dengan IPK memuaskan dan tingkat pendidikan pembimbing tugas akhirnya doktor. Dari kasus ini diperoleh nilai akurasi model yaitu sebesar 61,4%. Berdasarkan studi literatur yang telah dilakukan, artikel ini bertujuan untuk menganalisis faktor-faktor yang paling signifikan terhadap masa studi, membagi data ke dalam kelompok-kelompok yang lebih kecil berdasarkan hubungan antara variabel respon dan variabel penjelas.

2. METODE

CHAID ialah teknik eksplorasi nonparametrik yang menganalisis data berukuran besar, menguji setiap peubah penjelas yang dipakai dalam proses klasifikasi, serta menyusun kembali berdasarkan tingkat signifikan *Chi-square* terhadap peubah responnya (Gallagher, 2017). CHAID digunakan bila variabel *dependent* berskala nominal atau ordinal sehingga membentuk segmentasi yang membagi sampel menjadi beberapa sub-sub berbeda berdasarkan karakteristik yang akan diuji. Kejadian dilanjutkan membagi sub-sub tersebut menjadi sub yang lebih kecil berdasarkan peubah- penjelas dipakai.

Tabel 1. Keterangan data

No	Program Studi (S1)	Populasi (N=198)
1	Matematika	17
2	Fisika	27
3	Kimia	27
4	Biologi	33
5	Statistika	32
6	Informatika	28
7	Farmasi	34

Data yang digunakan diperoleh dari seluruh mahasiswa lulusan S1 FMIPA USK yang

lulus pada bulan 1 September 2020 hingga 31 Agustus 2022. Responden dalam penelitian ini berjumlah 198 mahasiswa yang telah menyelesaikan studi. Data ini diperoleh dari bagian Akademik FMIPA Universitas Syiah Kuala.

Peubah respon dalam penelitian ini adalah data kategorik berupa masa studi yang tergolong tiga kategori yaitu cepat, sedang, dan lama. Sedangkan peubah penjelas yang digunakan adalah sebagai berikut :

1. Jenis kelamin (X_1)
2. Asal sekolah (X_2)
3. Jurusan (X_3)
4. Tingkat pendidikan pembimbing (X_4)
5. IPK (X_5)

Peubah yang digunakan dalam penelitian ini dilihat pada Tabel 2 di bawah ini.

Tabel 2. Peubah penelitian yang digunakan

No	Peubah	Skala/Kategori
1	Masa studi (Y)	Ordinal 1 : Cepat (masa studi ≤ 8 semester) 2 : Sedang (8 semester < masa studi ≤ 9 semester) 3 : Lama (9 semester < masa studi ≤ 10 semester)
2	Jenis kelamin (X_1)	Nominal 1 : Perempuan 2 : Laki-laki
3	Asal sekolah (X_2)	Nominal 1 : Luar Banda Aceh 2 : Banda Aceh
4	Jurusan (X_3)	Nominal 1 : Matematika 2 : Fisika 3 : Kimia 4 : Biologi 5 : Informatika 6 : Farmasi 7 : Statistika
5	Tingkat pendidikan pembimbing (X_4)	Ordinal 1 : Bukan doktor 2 : Doktor
6	IPK (X_5)	Ordinal 1 : $\leq 2,75$ (Memuaskan) 2 : $2,76 - 3,50$ (Sangat memuaskan) 3 : $(\geq 3,51)$ (Pujian)

Langkah-langkah prosedur penelitian dalam penelitian ini adalah sebagai berikut:

1. Mempersiapkan data
 - a. Memisahkan jenis kelamin perempuan dan laki-laki menggunakan nama mahasiswa.
 - b. Menghitung masa studi dalam tahun dari Nomor Induk Mahasiswa (NIM) dan tanggal sidang.
2. Menganalisa secara deskriptif total data masa studi mahasiswa S1 FMIPA USK.
3. Menganalisa data menggunakan metode CHAID bertujuan untuk mendapatkan

faktor yang mempengaruhi masa studi mahasiswa S1 FMIPA USK dengan langkah sebagai berikut:

- a. Membuat peubah kategorik dari setiap peubah kontinu.
- b. Penggabungan kategorik dilakukan apabila peubah penjelasnya melebihi dua kategori. Dalam langkah penggabungan akan dilakukan uji *Chi-square* dan hipotesisnya:

H_0 : tidak terdapat hubungan antara dua peubah (saling bebas)

H_1 : terdapat hubungan antara dua peubah (tidak bebas)

Dengan statistik uji:

$$\chi^2 = \sum_i^r \sum_j^c \frac{(n_{ij} - \hat{u}_{ij})^2}{\hat{u}_{ij}}, \text{ dengan kriteria penolakan } H_0 \text{ jika}$$

$$\chi^2 > \chi^2_{(db, \alpha)}$$

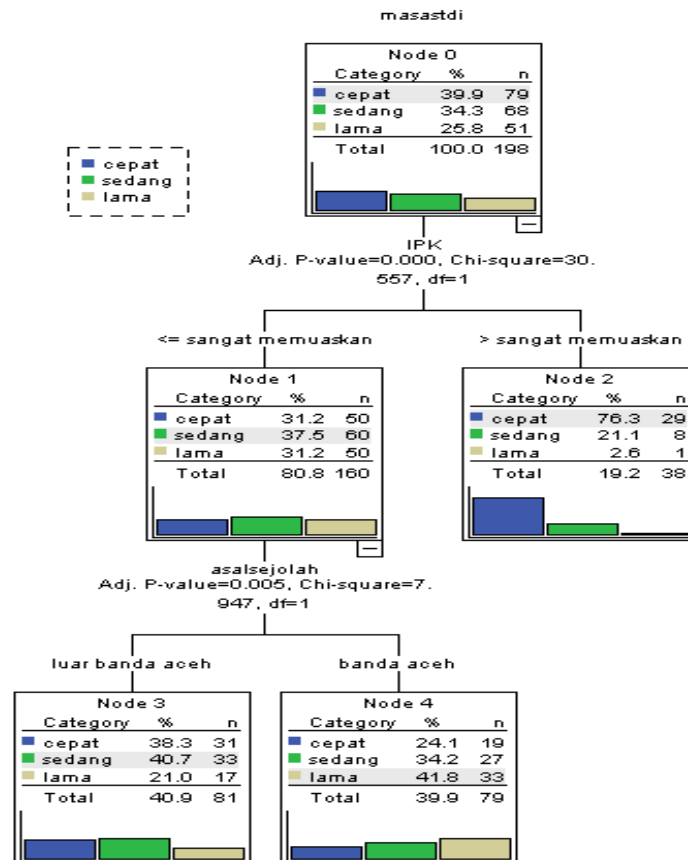
Jika uji yang dilakukan tidak signifikan, maka kategori tersebut digabungkan dan mengulangi langkah tersebut sampai tidak terdapat nilai kategori yang tidak signifikan. Jika uji yang dilakukan signifikan, maka tidak perlu digabungkan dan akan dihitung *adjusted p-value* terkecil dengan berbagai kategori untuk masing-masing peubah penjelas.

- c. Memilih penyekat peubah penjelas yang memiliki *adjusted p-value* terkecil. Jika *adjusted p-value* terkecil lebih besar dari beberapa nilai taraf nyata penyekat, maka penyekatan peubah penjelas dihentikan, dan simpul tersebut merupakan simpul terakhir. Sehingga, dapat dihasilkan sebuah diagram pohon.
 - d. Menghitung nilai akurasi menggunakan koefisien Kappa.
4. Interpretasi data, mengambil kesimpulan dengan menginterpretasikan diagram pohon CHAID.

3. HASIL DAN PEMBAHASAN

Data responden yang terkumpul dalam penelitian ini berjumlah 198 mahasiswa S1 FMIPA USK angkatan 2017 yang lulus dari tanggal 1 September 2020 hingga 31 Agustus 2022. Hasil analisa data dengan menggunakan metode CHAID pada tabel bagian *specification* terdapat 5 peubah penjelas, sedangkan pada variabel *result* hanya terdapat 2 peubah penjelas saja yang secara signifikan memberikan pengaruh pada model akhir. Peubah penjelas tersebut adalah IPK dan asal sekolah. Sedangkan peubah tingkat pendidikan pembimbing, jenis kelamin, dan jurusan tidak dimasukkan karena tidak signifikan berpengaruh terhadap model akhir. *Maximum tree depth* pada bagian *specification* adalah untuk membatasi pertumbuhan pohon klasifikasi. Pada tabel tersebut nilai *maximum tree dept* adalah 3 yaitu nilai batas yang sudah ditetapkan secara otomatis. Sedangkan *minimum cases in parent node* (jumlah minimum kasus pada *parent node*) sebanyak 100 dan *minimum cases in child node* (jumlah minimum kasus pada *child node*) sebanyak 5.

Hasil pembentukan model berupa pohon klasifikasi menggunakan software SPSS dengan node sebanyak 5 buah, node akhir (*terminal node*) sebanyak 2 buah dengan level kedalaman (*depth*) sebanyak 2. Peubah IPK dan asal sekolah merupakan peubah penjelas terbaik yang digunakan untuk membagi dan menerangkan peubah masa studi sebagai peubah respon. Diagram pohon yang dihasilkan dari analisis CHAID adalah sebagai berikut.



Gambar 1. Pohon Klasifikasi

Dari diagram pohon klasifikasi Gambar 1 dapat diinterpretasikan sebagai berikut:

- Terdapat 5 node yang dimulai dengan node 0 hingga node 4, dimana node 0 berisikan masa studi (cepat, sedang, dan lama), selanjutnya di node 1 dan node 2 berisikan IPK yang diperoleh yang dimana node 1 kategori (memuaskan) dan (sangat memuaskan) sedangkan di node 2 pujian, selanjutnya pada node 3 dan node 4 berisikan asal sekolah yang di mana node 3 (luar Banda Aceh), dan node 4 (Banda Aceh). Dengan jumlah terminal node sebanyak 3 yaitu node 2,3, dan 4 yang berarti terdapat 3 segmen dari masa studi mahasiswa S1 FMIPA USK.
- Peubah yang signifikan berpengaruh terhadap masa studi adalah IPK, dan asal sekolah.

Tabel 3. Nilai- *p* dan uji *chi-square* peubah penjelas dari diagram pohon

Peubah	Nilai- <i>p</i>	Nilai <i>Chi-square</i>
IPK	0,000	30,557
Asal Sekolah	0,005	7,947

Hipotesis pada pengujian *chi-square* adalah sebagai berikut.

H_0 : Tidak terdapat hubungan antara masa studi dengan IPK, dan asal sekolah.

H_1 : Terdapat hubungan antara masa studi dengan IPK, dan asal sekolah.

Dari tabel 4. dapat disimpulkan bahwa apabila dilakukan pengambilan keputusan berdasarkan nilai-*p*, di mana semua nilai tersebut kurang dari atau sama dengan $\alpha = 0,05$; yaitu 0,000; dan 0,005; maka H_0 ditolak, sehingga dapat disimpulkan bahwa terdapat hubungan antara masa studi dengan IPK dan asal sekolah.

- c. setiap terminal node dapat diartikan sebagai salah satu segmen dari masa studi mahasiswa S1 FMIPA USK.

Diagram pohon analisis CHAID pada gambar 1. menjelaskan bahwa masa studi mahasiswa S1 FMIPA USK yang lulus pada tanggal 1 September 2021 hingga 31 Agustus 2022 dapat dibagi menjadi beberapa segmen. Sedangkan untuk mengetahui peubah yang paling berpengaruh terhadap setiap kategori masa studi dapat dilihat dari persentase terbesar pada terminal node. Pada data tersebut terdapat 3 segmen yaitu sebagai berikut.

Tabel 4. Segmentasi dan persentase masa studi mahasiswa S1 FMIPA USK

Terminal node	Segmen	Karakteristik	Persentase mahasiswa
2	Ke-1 (node 0 dan 2)	Mahasiswa dengan masa studi cepat dengan IPK cumlaude.	76,3%
		Mahasiswa dengan masa studi sedang dan IPK cumlaude.	21,1%
		Mahasiswa dengan masa studi lama dan IPK cumlaude.	2,6%
3	Ke-2 (node 0,1, dan 3)	Mahasiswa dengan masa studi cepat dan IPK memuaskan atau sangat memuaskan dari asal sekolah luar Banda Aceh.	38,3%
		Mahasiswa dengan masa studi sedang dan IPK memuaskan atau sangat memuaskan dari asal sekolah luar Banda Aceh.	40,7%
		Mahasiswa dengan masa studi lama dan IPK memuaskan atau sangat memuaskan dari asal sekolah luar Banda Aceh.	21%
4	Ke-3 (node 0,1, dan 4)	Mahasiswa dengan masa studi cepat dan IPK memuaskan atau sangat memuaskan dari asal sekolah luar Banda Aceh.	24,1%
		Mahasiswa dengan masa studi sedang dan IPK memuaskan atau sangat memuaskan dari asal sekolah Banda Aceh.	34,2%
		Mahasiswa dengan masa studi lama dan IPK memuaskan atau sangat memuaskan dari asal sekolah Banda Aceh.	41,8%

Tabel 4 menginformasikan karakteristik per segmen dari kategori masa studi mahasiswa. Mahasiswa dengan masa studi cepat adalah 76,3 % yaitu persentase tertinggi segmen ke-2, yang berarti mahasiswa dengan masa studi cepat memiliki IPK *cumlaude*. Segmen yang masa studinya sedang memiliki persentase 40,7% adalah segmen ke-3, yaitu mahasiswa dengan masa studi sedang memiliki IPK memuaskan atau sangat memuaskan dari asal sekolah luar Banda Aceh. Sedangkan segmen yang masa studinya lama yaitu persentase tertinggi segmen ke-4, yang berarti mahasiswa dengan masa studi lama memiliki IPK memuaskan atau sangat memuaskan dari asal sekolah banda aceh dengan persentase 41,8%. Dari hasil metode CHAID tersebut telah di dapat segmen-segmen yang terbentuk dari data alumni mahasiswa FMIPA USK. Hasil klasifikasi ini bisa juga digunakan untuk mengetahui segmen-segmen pada mahasiswa yang masih aktif. Sudah diketahui dari data alumni bahwa segmen ke-4 memiliki masa studi yang lama, maka pada mahasiswa yang

masih aktif yang memiliki ciri-ciri seperti pada segmen tersebut perlu diwaspadai kalau mereka memiliki kecenderungan untuk terlambat dalam studinya atau memiliki masa studi yang lama.

Hasil yang diperoleh dari tabel risiko dan klasifikasi menyebutkan evaluasi dari seberapa bagus model bekerja dapat dilihat pada lampiran. Estimasi risiko sebesar 0,282 menunjukkan bahwa prediksi kategori oleh model (masa studi cepat, sedang atau lama) bagus untuk penelitian. Sehingga risiko dari klasifikasi yang salah untuk masa studi mahasiswa maksimal sebesar 28,2%. Tabel klasifikasi juga menunjukkan nilai yang sesuai, persentase keakuratan model mengklasifikasikan masa studi secara tepat sebesar 71,8%.

4. KESIMPULAN

Berdasarkan hasil analisis CHAID pada mahasiswa S1 FMIPA USK angkatan 2017 yang lulus pada tanggal 1 September 2020 sampai dengan 31 Agustus 2022 dapat disimpulkan bahwa :

1. Faktor-faktor prediksi yang signifikan memengaruhi masa studi mahasiswa secara berurutan adalah IPK dan asal sekolah.
2. Karakteristik kategori mahasiswa dengan masa studi cepat, adalah mahasiswa dengan masa studi cepat dan IPK *cumlaude*. Karakteristik kategori mahasiswa dengan masa studi sedang, adalah mahasiswa dengan masa studi sedang dan IPK memuaskan atau sangat memuaskan dari asal sekolah luar Banda Aceh. Karakteristik kategori mahasiswa dengan masa studi lama, adalah mahasiswa dengan masa studi lama dan IPK memuaskan atau sangat memuaskan dari asal sekolah Banda Aceh.
3. Persentase keakuratan model mengklasifikasi masa studi secara tepat adalah 71,8%.

REFERENSI

- Agresti, A. (2002). *Categorical data analysis second edition*. New York: John Wiley & Sons, LTD.
- Dudley, W. (1999). *Detecting and explicating interactions in categorical data*. Humboldt University of Berlin.
- Eherler, D. (2001). *Responder profiling with CHAID and dependency analysis*. Data Mining for Marketing Applications Workshop, University of Freiburg.
- Fitri, A. (2018). *Diajukan untuk memenuhi salah satu syarat guna mencapai gelar Sarjana dalam bidang ilmu Teknik Industri*. E Journal Analisis Waktu Kelulusan Mahasiswa Menggunakan Metode CHAID. Universitas Andalas Padang.
- FMIPA (2010). *Panduan Administrasi Akademik Program Sarjana dan Diploma Universitas Syiah Kuala*. http://fmipa.unsyiah.ac.id/images/download/panduan_akademik_unsyiah_2010p.pdf. Tanggal akses 29 Mei 2023.
- Jasma, A (2015). *Klasifikasi Masa Studi Mahasiswa S1 FMIPA Unsyiah Berdasarkan Karakteristik Akademik Mahasiswa*. Banda Aceh: Universitas Syiah Kuala.
- G.V.Kass. (1980). An Exploratory Technique for Investigating Large Quantities of Categorical Data. In *Applied Statistics*.
- Gallagher, C. (2000). *An iterative approach to classification analysis*. Journal of Applied Statistics.
- Kim, H., & Loh, W.-Y. (2009). Classification Trees with Unbiased Multiway Splits. *Journal of the American Statistical Association*.
- Loh, W. (1997). *Split selection methods for classification trees*. California: Wadsworth Belmonth.