

Identifikasi Faktor-Faktor yang Memengaruhi Prestasi Mahasiswa Menggunakan Regresi Logistik Ordinal dan *Random Forest Ordinal**

Studi Kasus Mahasiswa FMIPA IPB Angkatan 2015-2017

Zuhdiyah Izzatun Nisa¹, Agus M Soleh^{2‡}, and Hari Wijayanto³

^{1,2,3}Department of Statistics, IPB University, Indonesia

[‡]corresponding author: agusms@apps.ipb.ac.id

Copyright © 2021 Zuhdiyah Izzatun Nisa, Agus M Soleh, and Hari Wijayanto. This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Student achievement is the result of student learning processes and efforts. This research was conducted through a survey of students of the 2015-2017 FMIPA IPB with the selection of respondents using stratified random sampling. The purpose of this study is to identify the factors that influence the achievements of the 2015-2017 FMIPA IPB students using ordinal logistic regression and ordinal random forest. The response variable used is the PPKU GPA category and the last even semester GPA which is categorized based on the predicate of IPB graduation. The results of ordinal logistic regression get 7 explanatory variables that influence the PPKU GPA and 7 explanatory variables that influence the last even semester GPA. Explanatory variables that have a significant effect on ordinal logistic regression and become 10 variables with the highest level of importance in the ordinal random forest for both response variables are department, mother's education, internet access in a day for games, activity in the class, and active work on a group assignment.

Keywords: ordinal logistic regression, ordinal random forest, stratified random sampling, student achievement.

1 Pendahuluan

1.1 Latar Belakang

Prestasi adalah hasil yang didapatkan dari proses dan usaha belajar seseorang. Prestasi mahasiswa adalah hasil dari proses dan usaha yang telah dilakukan mahasiswa selama perkuliahan di perguruan tinggi. Prestasi mahasiswa dapat dipengaruhi oleh berbagai faktor. Menurut (Kuh *et al.*, 2006) faktor-faktor yang memengaruhi kemampuan akademik seorang mahasiswa dapat dibagi menjadi dua, yaitu pengalaman sebelum perkuliahan dan pengalaman selama perkuliahan.

Fakultas Matematika dan Ilmu Pengetahuan Alam Institut Pertanian Bogor (FMIPA IPB) merupakan fakultas terbesar yang dimiliki IPB. Program sarjana FMIPA IPB memiliki delapan departemen dengan karakteristik yang berbeda-beda. Perbedaan karakteristik tersebut dapat memengaruhi kondisi prestasi mahasiswanya dan salah satu ukuran yang bisa digunakan untuk mengukur prestasi mahasiswa adalah Indeks Prestasi Kumulatif (IPK).

Penelitian ini mengidentifikasi faktor-faktor yang memengaruhi prestasi mahasiswa FMIPA IPB angkatan 2015-2017 dengan pemilihan responden menggunakan penarikan contoh acak berlapis dua tahap. Lapisan yang digunakan dalam pemilihan responden adalah departemen dan angkatan. Pada penelitian sebelumnya, faktor-faktor yang memengaruhi IPK adalah lama belajar jika menggunakan regresi logistik biner, dan jika menggunakan regresi logistik multinomial 3 kategori adalah lama belajar, status pekerjaan ibu, pengeluaran uang per bulan, dan tambahan jam belajar (Fitriany, 2013).

Penelitian ini mengidentifikasi faktor-faktor yang memengaruhi prestasi mahasiswa menggunakan regresi logistik ordinal dan *random forest* ordinal. Kedua metode tersebut digunakan karena mampu mengidentifikasi peubah penjelas yang berpengaruh atau memiliki keterkaitan yang kuat terhadap peubah respon. Pengelompokan IPK didasarkan pada kriteria predikat kelulusan sesuai standar mutu IPB, yaitu mahasiswa yang berpredikat *Cum Laude* dengan kriteria $IPK > 3.5$ sebagai $Y = 3$, mahasiswa yang berpredikat sangat memuaskan dengan kriteria $3 < IPK \leq 3.5$ sebagai $Y = 2$, mahasiswa yang berpredikat memuaskan dengan kriteria $2.75 < IPK \leq 3$ sebagai $Y = 1$, dan dan mahasiswa yang lain tanpa predikat dengan kriteria $IPK \leq 2.75$ sebagai $Y = 0$ dengan asumsi ketentuan lain dalam menentukan predikat kelulusan diabaikan (IPB, 2018). Kemudian model yang terbentuk akan dievaluasi dan dibandingkan hasil serta kebaikan modelnya.

1.2 Tujuan

1. Mengidentifikasi karakteristik mahasiswa FMIPA IPB angkatan 2015-2017.
2. Mengidentifikasi faktor-faktor yang dapat memengaruhi IPK mahasiswa FMIPA IPB angkatan 2015-2017 menggunakan analisis regresi logistik ordinal dan *random forest*.
3. Membandingkan nilai evaluasi model regresi logistik ordinal dengan model *random forest* ordinal.

2 Metodologi

2.1 Proses Penarikan Contoh

Metode penarikan contoh yang digunakan dalam penelitian ini adalah penarikan contoh acak berlapis dua tahap. Penarikan contoh acak ini menggunakan departemen pada lapisan pertama dan angkatan pada lapisan kedua. Pemilihan lapisan ini karena diasumsikan setiap departemen dan angkatan memiliki karakteristik yang berbeda-beda. Populasi pada penarikan contoh ini adalah mahasiswa FMIPA IPB angkatan 2015-2017. Berikut tahapan dalam proses penarikan contohnya:

1. Menghitung nilai rata-rata dan ragam IPK mahasiswa di setiap departemen.
2. Menghitung jumlah contoh optimum untuk penarikan contoh acak berlapis menggunakan rumus (Scheaffer *et al.*, 1996)

$$n = \frac{\sum \frac{N_i^2 \sigma_i^2}{w_i}}{(\sum N_i \sigma_i^2) + N^2 \frac{B^2}{4}}$$

dengan batas toleransi (B) 5%.

3. Menghitung jumlah contoh di setiap lapisan departemen menggunakan rumus

$$n_i = n \left(\frac{N_i}{N} \right)$$

4. Menghitung jumlah contoh di setiap lapisan angkatan pada masing-masing departemen.
5. Melakukan penarikan contoh acak sederhana di setiap lapisan sesuai jumlah contoh yang dihitung sebelumnya

Jumlah contoh optimum yang didapatkan dari perhitungan rumus diatas adalah 269 contoh. Jumlah tersebut merupakan jumlah minimum responden yang harus diambil secara acak. Kemudian pada penelitian ini jumlah responden yang didapatkan sebanyak 300 orang dari 2014 jumlah mahasiswa FMIPA IPB angkatan 2015-2017.

2.2 Data

Data yang digunakan dalam penelitian ini adalah data primer dan data sekunder. Data primer diperoleh melalui survei terhadap mahasiswa FMIPA IPB angkatan 2015-2017 yang hasilnya akan digunakan sebagai peubah penjelas. Survei dilakukan melalui kuesioner yang dibagikan secara online dan offline pada tanggal 20 Februari 2019 sampai 17 Maret 2019.

Data sekunder diperoleh dari Direktorat Administrasi Pendidikan dan Penerimaan Mahasiswa Baru IPB berupa informasi nilai IPK mahasiswa FMIPA IPB yang selanjutnya dikategorikan berdasarkan predikat kelulusan dan digunakan sebagai peubah respon. Peubah respon yang digunakan pada penelitian ini adalah IPK PPKU dan IPK semester genap terakhir yang dikategorikan berdasarkan predikat kelulusan IPB. Pemilihan kedua peubah respon tersebut dengan asumsi bahwa karakteristik mahasiswa

tidak berubah selama tiga tahun terakhir. Terdapat 32 peubah penjelas yang digunakan pada penelitian ini. Pemilihan peubah penjelas tersebut didasarkan pada penelitian yang pernah dilakukan sebelumnya dengan beberapa penambahan. Tabel 1 menampilkan daftar peubah respon dan peubah penjelas yang digunakan dalam penelitian ini.

Tabel 1: Daftar Peubah Penjelas

Peubah	Nama Peubah	Skala
Y	IPK PPKU	Ordinal
X1	Departemen	Nominal
X2	Angkatan	Ordinal
X3	Jenis kelamin	Nominal
X4	Asal daerah	Nominal
X5	Riwayat penyakit	Nominal
X6	Jalur masuk IPB	Nominal
X7	Jumlah Saudara	Numerik
X8	Pendidikan ayah	Ordinal
X9	Pendidikan ibu	Ordinal
X11	Sumber biaya UKT	Nominal
X12	Pengeluaran per bulan	Numerik
X13	Memiliki penghasilan	Nominal
X14	Motivasi belajar dari dalam diri	Nominal
X15	Motivasi belajar dari luar diri	Nominal
X16	Teman dekat	Nominal
X17	Belajar kelompok	Nominal
X18	Gaya belajar	Nominal
X19	Waktu belajar mandiri	Nominal
X20	Media belajar utama	Nominal
X21	Akses internet untuk belajar	Ordinal
X22	Akses internet untuk media sosial	Ordinal
X23	Akses internet untuk <i>games</i>	Ordinal
X24	Akses internet untuk menonton film	Ordinal
X25	Akses internet untuk bekerja	Ordinal
X26	Waktu tidur	Ordinal
X27	Keaktifan di kelas	Ordinal
X28	Keaktifan mengerjakan tugas kelompok	Ordinal
X29	Keikutsertaan dalam perlombaan	Numerik
X30	Intensitas kunjungan ke perpustakaan	Ordinal
X31	Durasi waktu untuk organisasi	Numerik
X32	Intensitas datang tepat waktu	Ordinal
X32	Durasi perjalanan ke kampus	Ordinal

2.3 Prosedur Analisis Data

Analisis data pada penelitian ini menggunakan perangkat lunak R 3.5.2 dengan paket ordinal, MASS, caret, rpart, party, dan stats. Tahapan dalam analisis data yang akan dilakukan pada penelitian ini adalah sebagai berikut:

1. Mengumpulkan dan menginput data.
2. Melakukan analisis statistika deskriptif terhadap data.
3. Melakukan algoritme *stepwise* untuk mengatasi multikolinearitas (hubungan) pada peubah penjelas.
4. Membangun model regresi logistik ordinal menggunakan data keseluruhan peubah penjelas hasil algoritme *stepwise*, tahapannya:
 - (a) Menduga parameter regresi logistik ordinal.
 - (b) Melakukan pengujian parameter secara simultan untuk mengetahui peran seluruh peubah penjelas dalam model secara bersama.
 - (c) Melakukan pengujian parameter secara parsial untuk mengetahui peubah penjelas yang berpengaruh signifikan dalam model tersebut.
 - (d) Mencari nilai rasio odds dan melakukan interpretasi terhadap nilai rasio odds.
5. Mencari nilai ketepatan klasifikasi model regresi logistik ordinal.
6. Melakukan validasi silang pada model regresi logistik ordinal dengan menggunakan data latih sebesar 80% dan data uji sebesar 20% yang diulang sebanyak 100 kali.
7. Membentuk model *random forest* ordinal dengan pohon sebanyak 500 dan 6 peubah pemilah.
8. Melakukan evaluasi model *random forest* ordinal untuk data keseluruhan.
9. Melakukan validasi silang *random forest* ordinal dengan menggunakan data latih sebesar 80% dan data uji sebesar 20% yang diulang sebanyak 100 kali. Data latih juga digunakan untuk menentukan tingkat kepentingan peubah penjelas.
10. Membandingkan hasil evaluasi model regresi logistik ordinal dengan *random forest* ordinal.

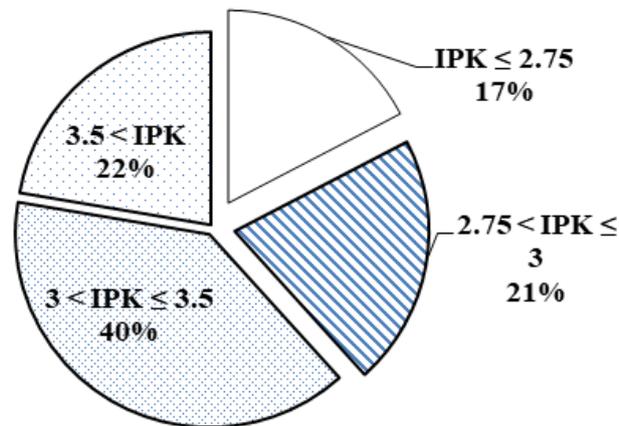
3 Hasil dan Pembahasan

3.1 Deskripsi Data

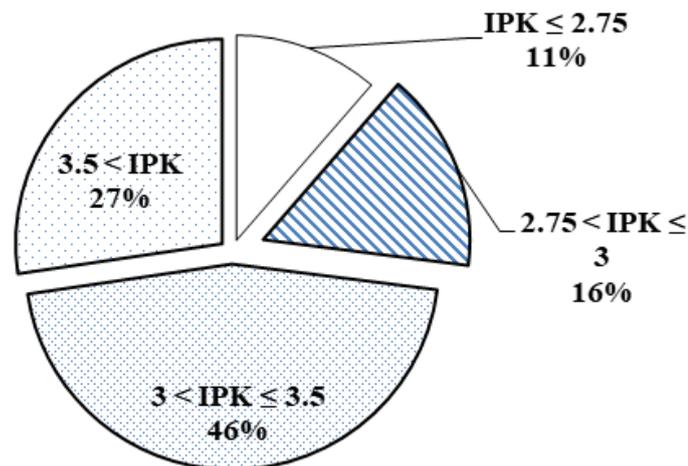
Data IPK PPKU mahasiswa FMIPA IPB angkatan 2015-2017 memiliki nilai rata-ran sebesar 3.21 dari 1913 mahasiswa. Nilai terendah, nilai median, dan nilai tertinggi dari data IPK PPKU yaitu sebesar 1.66, 3.27, dan 4. Data IPK semester genap terakhir mahasiswa FMIPA IPB angkatan 2015-2017 memiliki nilai rata-ran sebesar 3.13. Nilai terendah, nilai median, dan nilai tertinggi dari data IPK semester genap terakhir yaitu sebesar 1.65, 3.17, dan 4.

Gambar 1 menampilkan sebaran IPK PPKU responden yaitu IPK setelah melewati masa perkuliahan semester dua di IPB dan Gambar 2 menampilkan sebaran kategori IPK responden pada semester genap terakhir yang telah dilewati.

Responden pada penelitian ini diambil dari delapan departemen yang ada di FMIPA IPB yang menjadi lapisan pertama pada penarikan contoh acak berlapis. Gambar 3 menampilkan jumlah responden yang diambil pada setiap departemen. Untuk program studi aktuaria, responden yang diambil dikategorikan ke dalam departemen matematika.



Gambar 1: Sebaran IPK PPKU responden



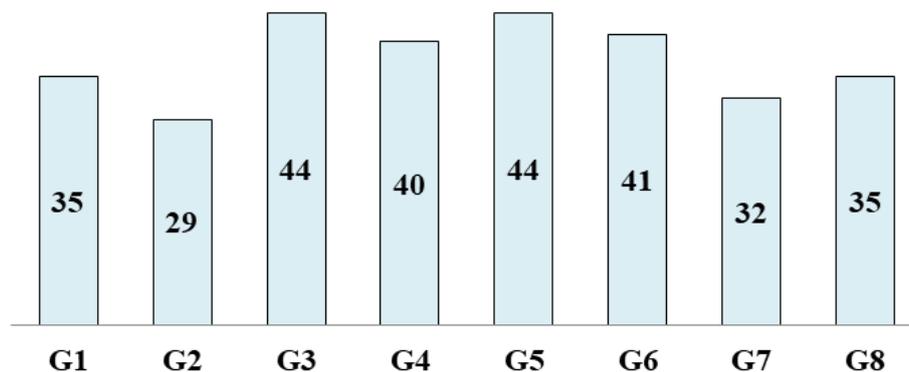
Gambar 2: Sebaran IPK semester genap terakhir responden

Lapisan kedua pada penarikan contoh acak berlapis adalah angkatan. Persentase jumlah responden yang diambil dari mahasiswa angkatan 2015 sebesar 32%, angkatan 2016 sebesar 33%, dan angkatan 2017 sebesar 35%. Data departemen dan angkatan selanjutnya digunakan sebagai peubah penjelas bersama 30 peubah penjelas lain yang datanya didapatkan melalui survei.

3.2 Regresi Logistik Ordinal

Analisis regresi logistik ordinal digunakan untuk mengidentifikasi peubah penjelas yang berpengaruh bagi prestasi mahasiswa, yang dalam penelitian ini adalah kategori IPK PPKU dan IPK semester genap terakhir. Asumsi yang harus dipenuhi dalam regresi logistik adalah tidak adanya multikolinearitas atau hubungan antar peubah penjelas. Pereduksian peubah menggunakan algoritme *stepwise* digunakan untuk menghasilkan peubah penjelas yang saling bebas.

Tujuh peubah penjelas yang ditampilkan pada Tabel 3 adalah departemen (X1), jalur masuk IPB (X6), pendidikan ibu (X9), motivasi belajar dari luar diri (X14), akses internet per hari untuk games (X22), keaktifan di kelas (X26), dan keaktifan



Gambar 3: Jumlah responden di setiap departemen

mengerjakan tugas kelompok (X27) pada taraf nyata 5%. Kemudian tujuh peubah penjelas tersebut digunakan untuk membentuk model regresi logistik ordinal yang baru.

Hasil pereduksian peubah penjelas menggunakan algoritme stepwise untuk peubah respon IPK semester genap terakhir (Y2) menghasilkan sembilan peubah penjelas yang saling bebas dari 32 peubah penjelas pada model awal. Kemudian dilakukan uji G untuk melihat ada atau tidaknya pengaruh peubah penjelas pada peubah respon. Nilai statistik uji G yang dihasilkan sebesar 308.36 dengan nilai-p 0.000 sehingga disimpulkan dari sembilan peubah penjelas hasil algoritme stepwise minimal ada satu peubah penjelas yang berpengaruh terhadap IPK semester genap terakhir pada taraf nyata 5%.

Selanjutnya dilakukan uji parsial untuk sembilan peubah penjelas tersebut. Hasil uji parsial menunjukkan terdapat tujuh peubah penjelas yang berpengaruh terhadap IPK semester genap terakhir. Tujuh peubah penjelas tersebut adalah departemen (X1), angkatan (X2), pendidikan ibu (X9), belajar kelompok (X16), akses internet per hari untuk games (X22), keaktifan di kelas (X26), dan keaktifan mengerjakan tugas kelompok (X27) pada taraf nyata 5%. Tabel 3 menampilkan tujuh peubah penjelas yang berpengaruh signifikan terhadap IPK semester genap terakhir.

Koefisien dari peubah penjelas menunjukkan perubahan nilai respon karena adanya perubahan nilai dari peubah penjelas (Hosmer *et al.*, 2013). Analisis regresi logistik ordinal memiliki peubah berskala non-metrik, sehingga interpretasi koefisien secara matematis sulit untuk dilakukan. Nilai rasio odds digunakan untuk mengetahui interpretasi dari setiap koefisien pada model regresi logistik ordinal.

Tabel 3 menampilkan nilai dugaan rasio odds untuk peubah respon IPK PPKU (Y1) dan Tabel 4 menampilkan nilai dugaan rasio odds untuk model yang terbentuk dari IPK semester genap terakhir (Y2). Uraian dari nilai rasio odds yang didapatkan adalah sebagai berikut:

1. Rasio odds peubah penjelas departemen (X1) memiliki nilai yang lebih kecil dari satu untuk setiap kategorinya. Departemen statistika menjadi kategori pembanding bagi departemen yang lainnya. Salah satu contoh interpretasinya adalah : untuk peubah respon IPK PPKU, nilai rasio odds departemen GFM adalah 0.16, artinya mahasiswa departemen GFM memiliki penurunan kemungkinan sebesar 84% untuk mendapatkan IPK PPKU yang lebih tinggi dibandingkan mahasiswa departemen statistika.

Tabel 2: Peubah penjelas yang berpengaruh terhadap IPK PPKU

Peubah	Keterangan	Koefisien	Nilai-p	OR
Intersep	1	-3.77		
	2	-2.54		
	3	-0.15		
X1	GFM	-1.87	0.00*	0.16
	Biologi	-1.60	0.00*	0.20
	Kimia	-1.75	0.00*	0.17
	Matematika	-1.25	0.01*	0.29
	Ilmu komputer	-1.43	0.00*	0.24
	Fisika	-2.24	0.00*	0.11
	Biokimia	-1.78	0.00*	0.17
X6	SBMPTN	-0.13	0.64	0.88
	UTMI	0.23	0.67	1.25
	BUD/Afirmasi	-1.89	0.00*	0.15
X9	Tamat SD	0.89	0.02*	2.42
	Tamat SMP	-0.65	0.08	0.52
	Tamat SMA	-0.31	0.35	0.73
X14	Perguruan tinggi	-0.32	0.32	0.73
	Tidak berpengaruh	0.85	0.01*	2.34
	<1 jam	-0.78	0.03*	0.46
X22	1-2 jam	0.15	0.62	1.17
	2-4 jam	-0.18	0.56	0.84
	>4 jam	-0.05	0.87	0.96
X26	Tidak aktif	-0.05	0.89	0.95
	Aktif	0.75	0.00*	2.12
X27	Aktif	0.77	0.03*	2.16
	Sangat aktif	0.27	0.22	1.32

Keterangan : *) signifikan pada taraf nyata 5%

2. Rasio odds peubah penjelas angkatan (X2) pada peubah respon IPK semester genap terakhir (Y2) berbeda untuk setiap kategorinya dan angkatan 2015 menjadi pembanding. Untuk mahasiswa angkatan 2016 rasio odds nya adalah 0.95, artinya mahasiswa angkatan 2016 memiliki penurunan kemungkinan sebesar 5% untuk mendapatkan IPK semester genap terakhir yang lebih tinggi dibandingkan mahasiswa angkatan 2015. Untuk mahasiswa angkatan 2017 rasio odds nya adalah 2.29, artinya mahasiswa angkatan 2017 memiliki kemungkinan 1.29 kali untuk mendapatkan IPK semester genap terakhir yang lebih tinggi dibandingkan mahasiswa angkatan 2015.
3. Rasio odds peubah penjelas jalur masuk IPB (X6) pada peubah respon IPK PPKU (Y1) berbeda untuk setiap kategorinya dan jalur masuk SNMPTN menjadi pembanding bagi jalur masuk yang lain. Salah satu contoh interpretasinya adalah : nilai rasio odds untuk jalur masuk BUD/Afirmasi sebesar 0.15, artinya mahasiswa yang masuk IPB melalui jalur BUD/Afirmasi memiliki penurunan kemungkinan sebesar 85% untuk mendapatkan IPK PPKU yang lebih tinggi dibandingkan jalur masuk SNMPTN.

Tabel 3: Peubah penjelas yang berpengaruh terhadap IPK semester genap terakhir

Peubah	Keterangan	Koefisien	Nilai-p	OR
Intersep	1	-2.17		
	2	-0.90		
	3	1.20		
X1	GFM	-1.35	0.01*	0.26
	Biologi	-1.90	0.05*	0.41
	Kimia	-1.16	0.01*	0.31
	Matematika	-1.82	0.06	0.44
	Ilmu komputer	-0.67	0.14	0.51
	Fisika	-1.79	0.00*	0.17
	Biokimia	-0.73	0.12	0.48
X2	2016	-0.05	0.86	0.95
	2017	0.83	0.00	2.29
X9	Tamat SD	0.77	0.05*	2.16
	Tamat SMP	-0.87	0.02*	0.42
	Tamat SMA	-0.14	0.66	0.87
X16	Perguruan tinggi	-0.27	0.38	0.76
	Tidak menyukai	0.78	0.00*	2.20
	<1 jam	-0.80	0.02*	0.45
	1-2 jam	0.16	0.59	1.17
X22	2-4 jam	0.19	0.52	1.21
	>4 jam	0.43	0.12	0.15
X26	Tidak aktif	-0.02	0.96	0.98
	Aktif	0.87	0.00*	2.38
X27	Aktif	0.94	0.01*	2.56
	Sangat aktif	0.22	0.31	1.25

Keterangan : *) signifikan pada taraf nyata 5%

- Rasio odds peubah penjelas pendidikan ibu (X9) memiliki nilai yang berbeda-beda di setiap kategorinya untuk kedua peubah respon dan kategori tidak tamat SD menjadi pembanding. Salah satu contoh interpretasinya adalah : rasio odds untuk kategori tamat SD adalah 2.34, artinya mahasiswa yang ibunya memiliki riwayat pendidikan hingga tamat SD memiliki kemungkinan 1.34 kali untuk mendapatkan IPK PPKU yang lebih tinggi dibandingkan mahasiswa yang ibunya tidak tamat SD.
- Rasio odds untuk peubah penjelas motivasi belajar dari luar diri (X14) adalah 2.34. Dapat diartikan bahwa mahasiswa yang motivasi belajarnya tidak dipengaruhi oleh faktor diluar dirinya sendiri memiliki kemungkinan sebesar 1.34 kali untuk IPK PPKU yang lebih tinggi dibandingkan mahasiswa yang motivasi belajarnya dipengaruhi oleh faktor diluar dirinya.
- Rasio odds untuk peubah penjelas suka atau tidaknya sistem belajar dengan berkelompok (X16) adalah 2.2. Artinya bahwa mahasiswa yang tidak menyukai sistem belajar kelompok memiliki kemungkinan 1.2 kali untuk mendapatkan IPK semester genap terakhir yang lebih tinggi dibandingkan dengan mahasiswa yang menyukai sistem belajar kelompok.
- Rasio odds untuk peubah penjelas akses internet per hari untuk games (X22)

memiliki nilai yang berbeda-beda di setiap kategorinya untuk kedua peubah respon an kategori 0 jam per hari menjadi pembanding. Salah satu contoh interpretasinya adalah : nilai rasio odds akses internet untuk bermain games kurang dari 1 jam per hari sebesar 0.45, artinya mahasiswa yang mengakses internet untuk bermain games kurang dari 1 jam per hari memiliki penurunan kemungkinan sebesar 55% untuk mendapatkan IPK PPKU yang lebih tinggi dibandingkan mahasiswa yang tidak mengakses internet untuk bermain games.

8. Rasio odds untuk peubah keaktifan di kelas (X26) memiliki nilai yang berbeda-beda di setiap kategorinya untuk kedua peubah respon. Salah satu contoh interpretasinya adalah : nilai rasio odds untuk mahasiswa yang aktif di kelas adalah 2.12, artinya mahasiswa yang aktif di kelas memiliki kemungkinan 1.12 kali untuk mendapatkan IPK PPKU yang lebih tinggi dibandingkan mahasiswa yang sangat tidak aktif ketika di kelas.
9. Rasio odds untuk peubah keaktifan mengerjakan tugas kelompok (X27) memiliki nilai yang lebih besar dari satu pada setiap ketegori untuk kedua peubah respon. Salah satu contoh interpretasinya adalah : nilai rasio odds untuk mahasiswa yang aktif mengerjakan tugas kelompok adalah 2.16, artinya mahasiswa yang aktif mengerjakan tugas kelompok memiliki kemungkinan 1.16 kali untuk mendapatkan IPK PPKU yang lebih tinggi dibandingkan mahasiswa yang tidak aktif mengerjakan tugas kelompok.

Evaluasi model pada penelitian ini menggunakan nilai akurasi, sensitivitas dan spesifisitas setiap kategori peubah respon. Penelitian ini juga menghitung nilai ketepatan klasifikasi menggunakan seluruh data untuk kedua peubah respon. Nilai ketepatan klasifikasi untuk peubah respon IPK PPKU (Y1) adalah 53.33% atau sebanyak 160 amatan tepat diprediksi oleh model dan yang lain tidak. Nilai ketepatan klasifikasi untuk peubah respon IPK semester genap terakhir (Y2) untuk seluruh data sebesar 43.67% atau sebanyak 131 amatan tepat diprediksi oleh model dan yang lain tidak. Tabel 4 menyajikan nilai evaluasi dari model regresi logistik ordinal untuk kedua peubah respon pada validasi silang dengan pengulangan sebanyak 100 kali. Nilai yang ditampilkan pada tabel 4 berupa nilai akurasi klasifikasi serta nilai sensitivitas dan nilai spesifisitas setiap kategori IPK.

Tabel 4: Nilai evaluasi model regresi logistik ordinal (%)

Nilai	Kategori IPK	Y1	Y2
Akurasi		48.18	39.46
Sensitivitas	0	15.92	32.51
	1	28.57	1.78
	2	83.07	69.79
	3	33.33	27.02
Spesifisitas	0	96.29	87.99
	1	99.91	98.16
	2	27.28	30.46
	3	88.29	90.97

Dapat dilihat nilai akurasi model IPK PPKU (Y1) pada validasi silang sebesar 48.18% dan model IPK semester genap terakhir (Y2) sebesar 39.46%. Hasil evaluasi model diatas

menunjukkan nilai sensitivitas terendah dimiliki oleh kategori satu. Hal ini dikarenakan data IPK yang masuk kategori dua lebih besar dan menjadi mayoritas untuk kedua peubah respon, sehingga kategori satu yang jumlahnya lebih sedikit cenderung diprediksi masuk ke kategori dua. Hal tersebut juga berdampak pada nilai spesifisitas terendah yang dimiliki oleh kategori dua.

3.3 *Random Forest Ordinal*

Pembentukan model pada klasifikasi random forest ordinal menggunakan 32 peubah penjelas yang diduga berpengaruh terhadap prestasi mahasiswa. Banyaknya pohon (k) yang digunakan untuk membangun model random forest ordinal pada penelitian ini adalah 500 pohon dan banyaknya peubah pemilah (m) yang digunakan adalah enam peubah. Sebelum ditentukan banyak pohon yang akan dibangun, dicoba terlebih dahulu pohon sebanyak 50, 100, 500 hingga 1000 pohon dan dihitung nilai rata-rata evaluasinya dengan validasi silang untuk menentukan banyak pohon optimal yang digunakan pada pemodelan random forest ordinal. Banyak peubah pemilah yang ditentukan sebelumnya juga dicobakan dengan 3, 6, dan 12 peubah, kemudian dihitung nilai rata-rata evaluasinya dengan menggunakan validasi silang untuk mendapatkan banyak peubah pemilah optimal. Validasi silang dilakukan dengan pengulangan sebanyak 100 kali dan diambil nilai rata-ratanya.

Tabel 5 menyajikan nilai evaluasi menggunakan validasi silang dari model random forest ordinal yang dibentuk untuk peubah respon IPK PPKU (Y1) dan IPK semester genap terakhir (Y2). Nilai yang ditampilkan merupakan nilai rata-rata dari pengulangan sebanyak 100 kali. Dapat dilihat nilai akurasi random forest ordinal untuk IPK PPKU (Y1) sebesar 45.57% dan untuk IPK semester genap terakhir (Y2) sebesar 38.76%.

Tabel 5 juga menampilkan nilai sensitivitas dan spesifisitas dari random forest ordinal. Nilai sensitivitas untuk setiap kategori pada kedua model sangat rendah, kecuali pada kategori dua. Hal ini disebabkan karena sebagian besar amatan diprediksi pada kategori yang mayoritas, yaitu kategori dua. Bahkan untuk model IPK PPKU nilai sensitivitas kategori nol dan satu adalah 0%, yang artinya tidak ada satupun amatan yang tepat diprediksi pada kategori nol dan satu. Kondisi ini menandakan model random forest ordinal kurang tepat untuk kasus ini karena hampir seluruh amatan diprediksi pada kelas mayoritas.

Tabel 5: Nilai evaluasi model *random forest* ordinal (dalam persentase)

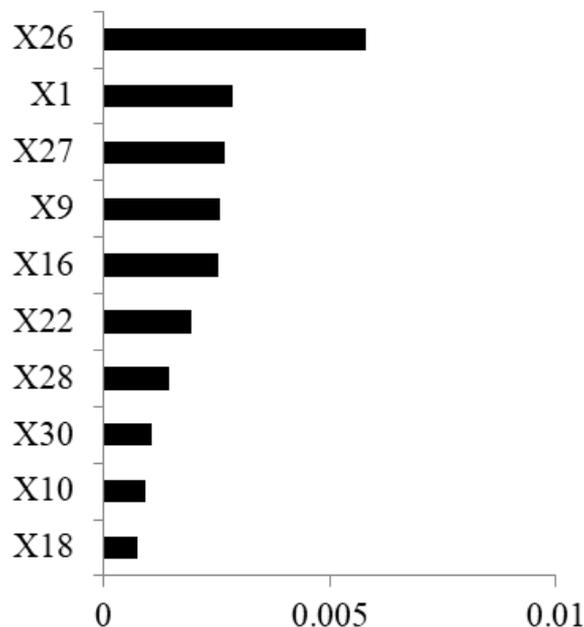
Nilai	Kategori IPK	Y1	Y2
Akurasi		45.57	38.76
Sensitivitas	0	0.00	0.17
	1	0.00	0.00
	2	99.34	98.36
	3	0.99	2.02
	1	100.00	99.92
Spesifisitas	2	100.00	99.96
	3	1.25	1.54
	4	98.99	98.68

Ukuran kepentingan peubah penjelas yang digunakan pada penelitian ini adalah VIMS-RPS, karena sesuai dengan penelitian yang telah dilakukan sebelumnya

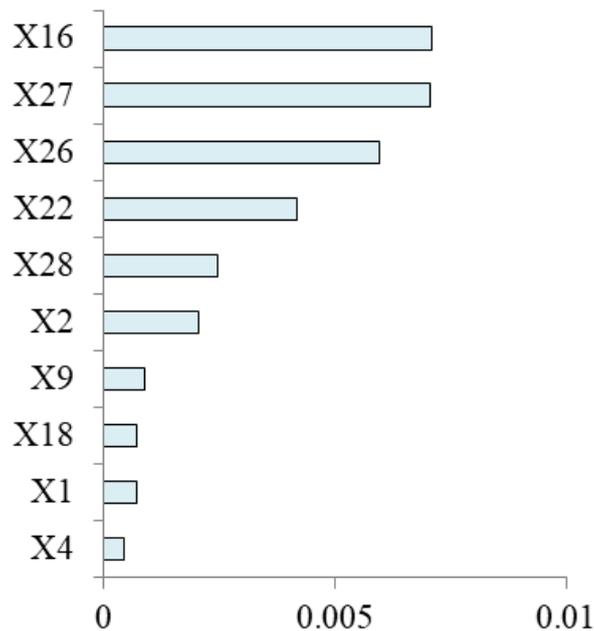
pengukuran peubah penciri yang terbaik untuk respon ordinal adalah VIMs-RPS (Janitza *et al.*, 2014). Gambar 7 menampilkan ilustrasi sepuluh peubah penjelas yang memiliki tingkat kepentingan peubah tertinggi dengan 500 pohon yang dibentuk dan enam peubah pemilah.

Gambar 4 menampilkan tingkat kepentingan peubah penjelas bagi peubah respon IPK PPKU (Y1) menggunakan VIMs-RPS. Dapat dilihat terdapat sepuluh peubah penjelas yang memiliki nilai kepentingan peubah penjelas lebih tinggi dari peubah penjelas yang lain. Peubah penjelas tersebut adalah keaktifan di kelas (X26), departemen (X1), pendidikan ibu (X9), keaktifan mengerjakan tugas kelompok (X27), belajar kelompok (X16), akses internet per hari untuk games (X22), keikutsertaan dalam lomba/kompetisi (X28), durasi waktu untuk organisasi/kepanitiaan per pekan (X30), sumber biaya UKT (X10), dan waktu belajar mandiri (X18).

Gambar 5 menampilkan tingkat kepentingan peubah penjelas bagi peubah respon IPK semester genap terakhir (Y2) menggunakan VIMs-RPS. Dapat dilihat terdapat sepuluh peubah penjelas yang memiliki nilai kepentingan peubah penjelas lebih tinggi dari peubah penjelas yang lain. Peubah penjelas tersebut adalah belajar kelompok (X16), keaktifan mengerjakan tugas kelompok (X27), keaktifan di kelas (X26), akses internet per hari untuk games (X22), keikutsertaan dalam lomba/kompetisi (X28), angkatan (X2), waktu belajar mandiri (X18), departemen (X1), dan asal daerah (X4).



Gambar 4: Tingkat kepentingan peubah penjelas pada Y1



Gambar 5: Tingkat kepentingan peubah penjelas pada Y2

4 Simpulan

Identifikasi karakteristik mahasiswa FMIPA IPB angkatan 2015-2017 memberikan hasil bahwa mahasiswa FMIPA IPB memiliki karakteristik yang berbeda-beda di setiap departemennya, namun cenderung sama untuk setiap angkatan. Model terbaik yang didapatkan dari penelitian ini adalah model regresi logistik ordinal karena menghasilkan nilai akurasi yang lebih tinggi serta nilai sensitivitas dan spesifisitas yang lebih seimbang di setiap kategori IPK dibandingkan model random forest ordinal. Peubah penjelas yang berpengaruh signifikan terhadap IPK PPKU dan masuk kedalam 10 peubah penjelas dengan tingkat kepentingan tertinggi pada random forest ordinal adalah departemen (X1), pendidikan ibu (X9), akses internet per hari untuk games (X22), keaktifan di kelas (X26) dan keaktifan mengerjakan tugas kelompok (X27). Peubah penjelas berpengaruh signifikan terhadap IPK semester genap terakhir dan masuk kedalam 10 peubah penjelas dengan tingkat kepentingan tertinggi pada random forest ordinal adalah departemen (X1), angkatan (X2), pendidikan ibu (X9), belajar kelompok (X16), akses internet per hari untuk games (X22), keaktifan di kelas (X26), dan keaktifan mengerjakan tugas kelompok (X27). Dapat disimpulkan terdapat lima peubah penjelas yang berpengaruh pada kedua peubah respon dengan kedua metode yaitu departemen, pendidikan ibu, akses internet per hari untuk games, keaktifan di kelas, dan keaktifan mengerjakan tugas kelompok.

Daftar Pustaka

Fitriany, N. (2013). *Identifikasi Faktor-Faktor yang Mempengaruhi Indeks Prestasi Kumulatif Menggunakan Regresi Logistik Biner dan Multinomial[skripsi]*. Institut Pertanian Bogor: Bogor(ID).

Hosmer, D., Lemeshow, S., & Sturdivan, R. (2013). *Applied Logistic Regression*. New Jersey (US): Wiley.

IPB, I. P. B. (2018). *Buku Panduan Pendidikan Sarjana*. Bogor (ID): IPB Press.

Janitza, S., Tutz, G., & Boulesteix, A.-L. (2014). Random forests for ordinal response data: prediction and variable selection. Technical report.

Kuh, G., Kinzie, J., Buckley, J., Bridges, B., & JC, H. (2006). What matters to student success: a review of the literature. Technical report.

Scheaffer, R., Mendenhall, W., & Ott, L. (1996). *Elementary Survey Sampling*. Boston (US): Brooks/Cole.