

POHON KLASIFIKASI DAN POHON REGRESI KEBERHASILAN MAHASISWA PASCASARJANA PROGRAM STUDI STATISTIKA IPB

Ida Mariati H.¹⁾, Khairil A. Notodiputro²⁾, dan Bambang Juanda²⁾

¹⁾ Jurusan Matematika, FMIPA UNCEN

²⁾ Departemen Statistika, FMIPA IPB

Ringkasan

Selama tiga tahun terakhir, untuk dua semester pertama, selalu terdapat mahasiswa Program Studi Statistika Sekolah Pascasarjana IPB yang tidak sanggup memenuhi persyaratan yang telah ditetapkan. Karena itu menjadi penting untuk mengetahui faktor-faktor yang mempengaruhi keberhasilan mahasiswa dengan menelusuri latar belakang pendidikan S1 dan prestasi akademik mahasiswa selama S1. Penelitian ini dilakukan pada mahasiswa S2 Program Studi Statistika angkatan 1991 - 2002. Pendekatan analisis yang digunakan adalah Metode Pohon Regresi (Regression Tree).

Faktor-faktor yang diduga mempengaruhi keberhasilan mahasiswa adalah jenis kelamin, status perkawinan, usia, pekerjaan, asal perguruan tinggi S1, program studi S1, nilai mutu rata-rata (NMR) yang dicapai untuk semua mata kuliah selama S1, nilai mutu rata-rata (NMR) untuk semua mata kuliah statistika selama S1, dan nilai mutu rata-rata (NMR) untuk semua mata kuliah matematika selama S1. Hasil analisis dari ke tiga pohon regresi menunjukkan bahwa jenis pekerjaan, asal perguruan tinggi dan NMR STK S1 adalah peubah yang sangat berpengaruh terhadap keberhasilan mahasiswa S2 Statistika. Mahasiswa yang berasal dari perguruan tinggi BHMN lebih unggul dibandingkan mahasiswa yang berasal dari perguruan tinggi Jawa non BHMN dan Luar Jawa.

Kata kunci : *NMR, statistika deskriptif, pohon klasifikasi, pohon regresi*

PENDAHULUAN

Institut Pertanian Bogor merupakan salah satu universitas terkemuka di Indonesia yang menyelenggarakan pendidikan pascasarjana sejak tahun 1975. Dari hanya tujuh program studi yang dibuka tahun itu, salah satunya adalah program studi Statistika Terapan, sekarang berkembang menjadi 35 program studi Magister Sains (S2). Program-program studi tersebut dirancang secara khusus untuk menumbuhkan dan meningkatkan kemampuan para mahasiswa untuk berprestasi lebih baik dalam hidupnya, baik untuk pribadinya, keluarga dan masyarakat, negara, serta untuk sesama umat manusia (IPB, 2001).

Penerimaan mahasiswa baru di Sekolah Pascasarjana IPB dilakukan berdasarkan penilaian transkrip dan indikator lain seperti akreditasi, status perguruan tinggi dan kesesuaian bidang studi selama pendidikan sarjana dan atau pasca sarjana. Kelemahan penilaian transkrip ini adalah sangat beragamnya kualitas lulusan sarjana meskipun nilai transkripnya sama. Hal ini disebabkan karena fasilitas pendidikan di

perguruan tinggi yang ada di Indonesia belum merata dan memadai terutama dosen, laboratorium dan buku-buku di perpustakaan.

Calon mahasiswa yang masuk ke program studi Statistika adalah lulusan S1 Statistika/Matematika, atau kependidikan yang berlatar belakang Matematika dan Statistika dari berbagai Perguruan Tinggi yang ada di Indonesia, dan ada juga yang bukan lulusan Statistika/Matematika. Bagi mahasiswa yang memiliki latar belakang yang kurang di bidang Matematika dan Statistika diharuskan mengambil beberapa mata kuliah program studi S1 yang dianggap perlu sebagai prasyarat di semester I (satu).

Seorang mahasiswa pada program Magister Sains harus mempertahankan NMR (Nilai Mutu Rata-rata) tidak kurang dari 3,00 pada seluruh mata kuliah yang ditempuhnya. Apabila pada akhir semester II (dua) seorang mahasiswa tidak dapat mempertahankan NMR tersebut di atas, tidak diizinkan meneruskan pendidikannya di SPs-IPB. Untuk mahasiswa yang diterima dengan status percobaan, NMR tersebut diatas

harus dapat dipertahankan pada akhir semester I. Jika tidak, maka tidak diizinkan meneruskan pendidikannya di SPs - IPB .

Pada program studi Statistika, selama 3 tahun terakhir ini, untuk dua semester pertama selalu terdapat mahasiswa yang tidak sanggup memenuhi persyaratan yang telah ditetapkan oleh sekolah pascasarjana (SPs) seperti yang terlihat pada Tabel 1 di bawah ini.

Tabel 1. NMR Mahasiswa Program Studi Statistika yang Tidak Memenuhi Persyaratan

Tahun. Masuk	Mahasiswa	Indeks Prestasi		NMR
		Sem I	Sem II	
2000	1	2,50	-	2,50 (keluar)
	2	2,83	2,50	2,67
	3	2,83	3,00	2,92
	4	3,17	2,75	2,96
	5	2,50	2,50	2,50
2001	1	3,33	2,63	2,98
	2	3,00	2,63	2,82
	3	2,50	2,00	2,25 (keluar)
	4	2,33	-	2,33 (pindah PS)
	5	2,35	2,88	2,62 (pindah PS)
	6	2,50	-	2,50 (keluar)
2002	1	3,00	2,88	2,94
	2	2,67	-	2,67 (keluar)
	3	2,00	-	2,00 (pindah PS)
	4	2,83	-	2,83 (keluar)
	5	2,83	2,83	2,83
	6	3,00	2,83	2,92
	7	2,00	-	2,00 (keluar)

Sumber : Divisi Akademik Sekolah Pascasarjana IPB

Dari Tabel 1 terlihat bahwa untuk angkatan 2000 yang tidak memenuhi persyaratan sebesar 25% dari 20 orang, angkatan 2001 sebesar 30% dari 20 orang dan untuk angkatan 2002 yang tidak memenuhi persyaratan sebesar 31% dari 22 orang.

Persyaratan umum yang diterapkan tersebut menuntut kualifikasi mahasiswa yang unggul di dalam proses pendidikan selama S2. Dalam hal ini diperlukan pemilihan mahasiswa yang tepat, sehingga menjadi penting untuk mengetahui faktor-faktor yang mempengaruhi keberhasilan mahasiswa, yaitu dengan menelusuri latar belakang pendidikan S1 dan prestasi akademik mahasiswa selama di S1.

Tujuan dan Manfaat Penelitian

Penelitian ini bertujuan untuk mengetahui faktor-faktor penciri yang berpengaruh terhadap keberhasilan mahasiswa program studi Statistika Sekolah Pascasarjana IPB, dengan cara membuat model pohon regresi. Manfaat dari hasil penelitian ini diharapkan dapat memberikan gambaran atau petunjuk bagi langkah selanjutnya dalam penyeleksian calon mahasiswa baru di Sekolah

Pascasarjana IPB khususnya program studi Statistika guna lebih meningkatkan mutu dan kualitas lulusan, serta untuk kelancaran studi di IPB.

DATA DAN METODE

Data

Data yang digunakan dalam penelitian ini adalah data mahasiswa S2 Program Studi Statistika IPB angkatan 1991 sampai dengan 2002 yaitu berjumlah 148 mahasiswa. Dari jumlah ini 124 (83,78%) berhasil tamat dan 24 (16,22 %) dikeluarkan (dropt out). Mahasiswa yang tertera pada Tabel 1 diikutsertakan dalam analisis yaitu termasuk mahasiswa yang tidak berhasil. Bagi mahasiswa yang tidak dikeluarkan sampai saat ini masih ada yang aktif kuliah sebanyak 6 orang dan yang sudah berhasil tamat 3 orang. Data diperoleh dari Divisi Akademik Sekolah Pascasarjana IPB. Proses pengambilan data berlangsung selama 4 bulan.

Ada 3 (tiga) peubah respon yang digunakan yaitu Nilai Mutu Rata-rata (NMR) S2, masa studi S2 dan peringkat kelas yang dilihat dari Indikator Kinerja (IK). Indikator kinerja diperoleh dari selisih Nilai Mutu Rata-rata (NMR) S2 dengan masa studi yang ditempuh setelah masing-masing terlebih dahulu distandarisasi dengan cara mengurangi nilai indikator tersebut dengan rata-ratanya kemudian dibagi dengan simpangan bakunya.

$$IK = Z_{nmr} - Z_{M.Studi} \quad \text{dimana } Z = \frac{X - \bar{X}}{\sigma}$$

Kemudian hasil indikator kinerja tersebut dibagi menjadi 3 kelompok selang di dalam pembagian peringkat kelas.

Kelompok peringkat kelas adalah :

1. Kelas 1 (baik), jika $IK \geq 1$
2. Kelas 2 (cukup), jika $-1 < IK < 1$
3. Kelas 3 (kurang), jika $IK \leq -1$
4. Kelas 4 (Gagal / Do)

Sedangkan peubah penjelasnya, antara lain :

1. Jenis Kelamin (JK = 1, untuk laki-laki dan 0, untuk perempuan)
2. Status Perkawinan (Status = 1, untuk kawin 0, untuk belum kawin)
3. Usia. Usia pada saat masuk sekolah pascasarjana dalam tahun
4. Pekerjaan Pekerj = 3, untuk lainnya (bukan dosen maupun peneliti) 2, untuk Peneliti, dan 1, untuk Dosen
5. Asal P.Tinggi S1
 Asa_PT = 2, untuk BHMN
 1, untuk Jawa non BHMN

- 0, untuk Luar Jawa
- 6. Program Studi S1
 - P_Studi = 3, untuk Statistika
 - 2, untuk Matematika
 - 1, untuk Pend. Matematika
 - 0, untuk lainnya
- 7. Nilai Mutu Rata-rata (NMR) yang dicapai untuk semua mata kuliah selama S1.
- 8. Nilai Mutu Rata-rata (NMR) untuk semua Statistika selama S1 mata kuliah.
- 9. Nilai Mutu Rata-rata (NMR) untuk semua mata kuliah Matematika selama S1.

Metode Analisis

Analisis data dilakukan dalam dua tahap. Tahap pertama adalah menentukan pohon optimal dalam CART yang diperoleh melalui proses pemangkasan (pruning). Kemudian dilakukan uji tingkat ketepatan pengklasifikasian amatan dari pohon regresi yang terbentuk dengan menggunakan semua data yang dianalisis. Disamping itu juga dilakukan validasi model untuk menunjukkan konsistensi dari pohon yang dihasilkan dengan menggunakan data baru.

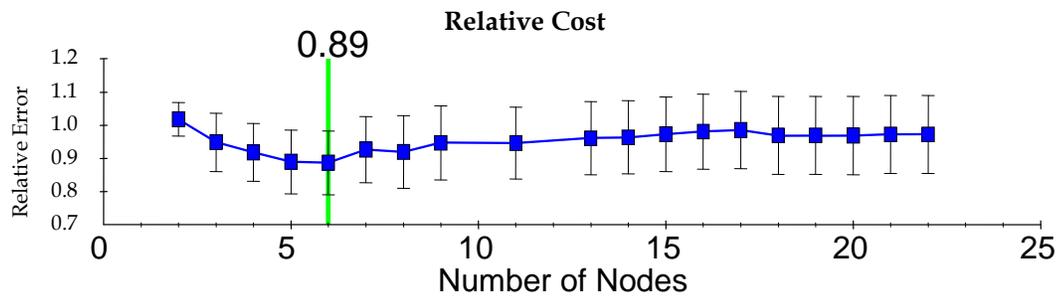
Tahap kedua adalah menentukan nilai dugaan respon bagi setiap simpul akhir. Proses analisis dilakukan dengan menggunakan software Minitab versi 13.20 dan software CART versi 4.0.

HASIL DAN PEMBAHASAN

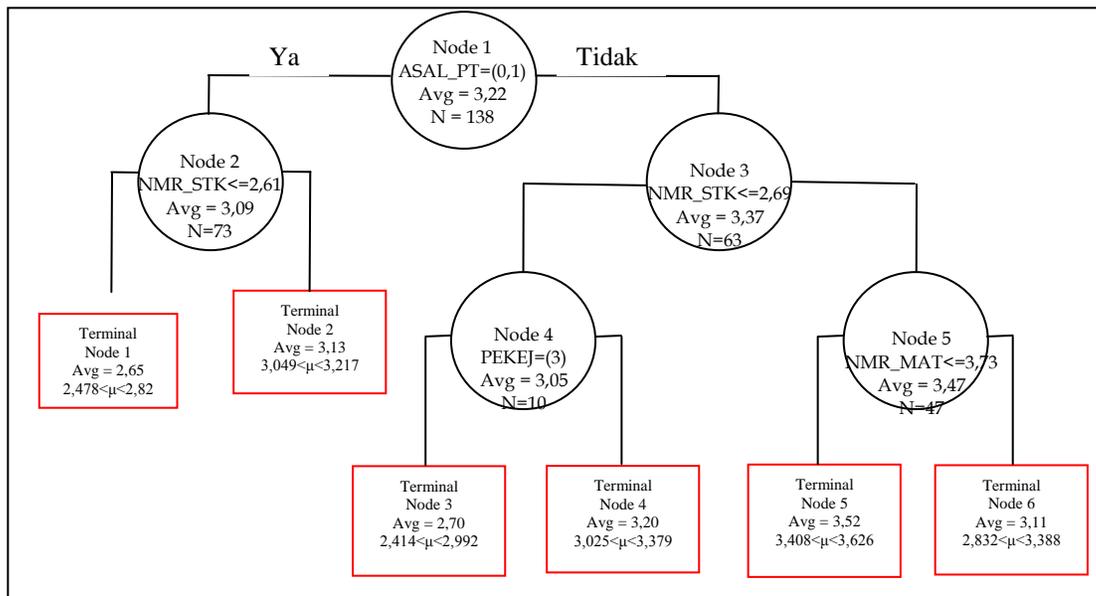
Pohon Regresi Untuk Peubah Respon NMR S2

Pada tahap awal dihasilkan pohon maksimal berukuran 22 simpul terminal. Dan melalui pemangkasan berdasarkan aturan *Cost Complexity Minimum* dan penggunaan validasi Silang lipat-10 (*10-fold Cross Validation*) dengan nilai R^{CV} (*x-val Relative Error*) terkecil sebesar 0,89, diperoleh pohon optimum (Gambar 1).

Pohon ini mengandung enam buah simpul terminal dengan tingkat ketepatan pengklasifikasian amatan sebesar 72,46% (Gambar 2). Hasil validasi model dengan menggunakan data baru sebanyak 10 mahasiswa diperoleh tingkat ketepatan 60% dan 40% berada di atas selang.



Gambar 1. Plot R^{CV} (*x-val Relative Error*) Peubah Respon NMR S2



Gambar 2. Pohon regresi Optimum Peubah Respon NMR S2

Peubah penjelas yang pertama menyekat adalah asal perguruan tinggi S1. Sehingga peubah ini merupakan peubah dominan. Peubah yang lain yang muncul adalah peubah NMR S1, NMR STK, NMR MAT, dan pekerjaan.

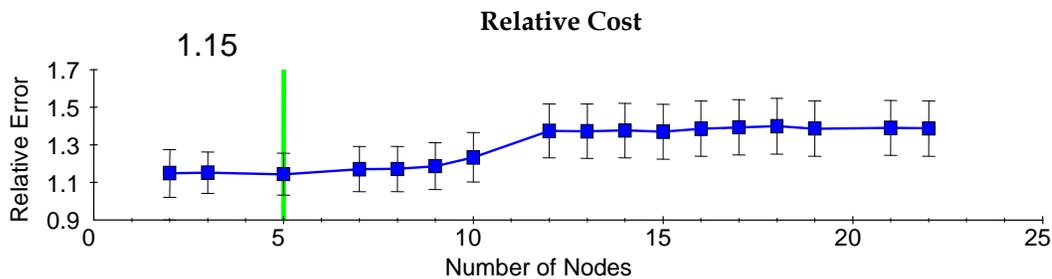
Terlihat bahwa kelompok mahasiswa yang berasal dari Perguruan Tinggi BHMN, $NMR S1 > 2,69$, $NMR MAT \leq 3,73$ merupakan kelompok dengan NMR S2 tertinggi, sedangkan untuk kelompok mahasiswa yang berasal dari Perguruan Tinggi BHMN, $NMR S1 \leq 2,69$ serta pekerjaan bukan dosen maupun peneliti dan kelompok mahasiswa yang berasal dari Perguruan Tinggi Jawa Non BHMN atau Luar Jawa, $NMR STK \leq 2,61$ merupakan kelompok mahasiswa yang diprediksikan akan mengalami kegagalan (DO). Pada simpul akhir 4 dan simpul akhir 5 terlihat bahwa mahasiswa yang memiliki $NMR MAT \leq 3,73$ memiliki rata-rata NMR S2 lebih tinggi dibandingkan mahasiswa yang memiliki $NMR MAT > 3,73$. Kesimpulan ini terlihat agak aneh, hal ini terjadi karena mahasiswa yang memiliki NMR

$MAT > 3,73$ ternyata sebagian besar berasal dari program studi selain statistika dan matematika, di mana satuan kredit yang diambil mata ajaran matematika tidak lebih dari 10% saja.

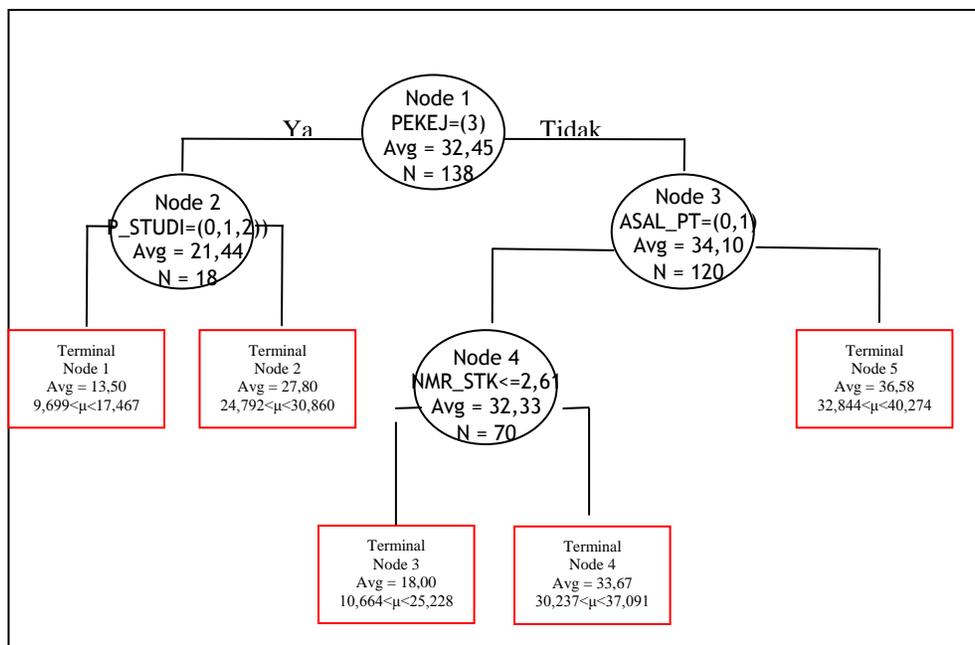
Pohon Regresi Untuk Peubah Respon Masa Studi S2

Pada tahap awal dihasilkan pohon maksimal berukuran 22 simpul terminal.

Kemudian melalui pemangkasan berdasarkan aturan *Cost Complexity Minimum* dan penggunaan validasi Silang lipat-10 (*10-fold Cross Validation*) dengan nilai R^{CV} (*x-val Relative Error*) terkecil sebesar 1,15, diperoleh pohon optimum (Gambar 3). Pohon ini mengandung lima buah simpul terminal dengan tingkat ketepatan pengklasifikasian amatan sebesar 65,21% (Gambar 4). Hasil validasi model dengan menggunakan data baru sebanyak 10 mahasiswa diperoleh tingkat ketepatan 50%, 20% berada di bawah selang dan 40% berada di atas selang.



Gambar 3. Plot R^{CV} (x-val Relative Error) Peubah Respon Masa Studi S2



Gambar 4. Pohon Regresi Optimum Peubah Respon Masa Studi S2

Peubah penjelas yang pertama menyekat adalah jenis pekerjaan. Sehingga peubah ini merupakan peubah dominan. Peubah yang lain yang muncul adalah peubah asal Program studi S1, Asal Perguruan Tinggi dan NMR STK.

Terlihat bahwa kelompok mahasiswa yang bukan sebagai dosen maupun peneliti, dan berasal dari program studi Matematika, Pendidikan Matematika dan Program studi lainnya serta kelompok mahasiswa yang bekerja sebagai dosen maupun peneliti, berasal dari Perguruan Tinggi Jawa non BHMN atau Luar Jawa serta $NMR_STK \leq 2,61$ merupakan kelompok mahasiswa yang diprediksikan akan mengalami kegagalan (DO).

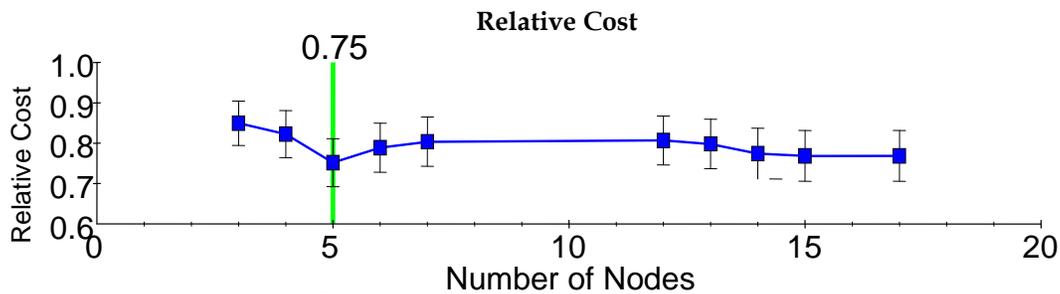
Kemudian melalui pemangkasan berdasarkan aturan *Cost Complexity Minimum* dan penggunaan validasi Silang lipat-10 (*10-fold Cross Validation*) dengan nilai R^{CV} (*x-val Relative Error*) terkecil sebesar 0,75, diperoleh pohon optimum (Gambar 5). Pohon ini mengandung lima buah simpul terminal dengan tingkat ketepatan pengklasifikasian amatan sebesar 62,32% (Gambar 6). Hasil validasi model dengan menggunakan data baru sebanyak 10 mahasiswa diperoleh tingkat ketepatan 50%.

Peubah penjelas yang pertama menyekat adalah jenis pekerjaan. Sehingga peubah ini merupakan peubah dominan. Peubah yang lain yang muncul adalah Asal Perguruan Tinggi, NMR S1 dan NMR STK.

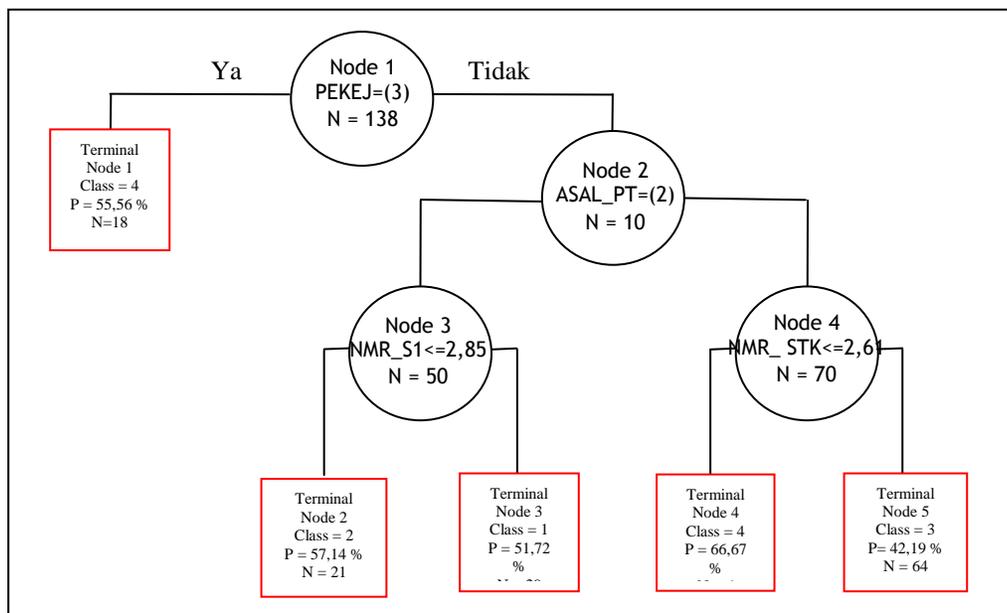
Terlihat bahwa kelompok mahasiswa sebagai dosen maupun peneliti, berasal dari Perguruan Tinggi BHMN dan $NMR_S1 > 2,85$ merupakan kelompok mahasiswa yang kelas 1 (baik).

Pohon Regresi Untuk Peubah Respon Peringkat Kelas

Pada tahap awal dihasilkan pohon maksimal berukuran 17 simpul terminal.



Gambar 5. Plot R^{CV} (*x-val Relative Error*) Peubah Respon Peringkat Kelas



Gambar 6. Pohon Regresi Optimum Peubah Respon Peringkat Kelas

Dari ketiga pohon regresi di atas, terlihat bahwa jenis pekerjaan, asal perguruan tinggi dan NMR STK muncul sebagai peubah yang diduga berpengaruh terhadap keberhasilan mahasiswa S2 Statistika SPs IPB. Mahasiswa yang bukan dosen dan peneliti kemungkinannya lebih besar mengalami kegagalan dalam menempuh studi S2 di program studi Statistika IPB. Mahasiswa yang berasal dari perguruan tinggi BHMN lebih unggul dibandingkan mahasiswa yang berasal dari perguruan tinggi Jawa non BHMN dan Luar Jawa.

KESIMPULAN DAN SARAN

Kesimpulan

Kajian ini menunjukkan bahwa keberhasilan mahasiswa statistika IPB dipengaruhi oleh status perguruan tinggi, NMR pendidikan S1, NMR mata kuliah Matematika selama pendidikan S1, serta jenis pekerjaan mahasiswa. Secara lebih rinci keberhasilan mahasiswa statistika dapat dideskripsikan sebagai berikut :

1. Kelompok mahasiswa dengan NMR S2 tertinggi yaitu mahasiswa yang berasal dari Perguruan Tinggi BHMN, $NMR S1 > 2,69$, dan $NMR MAT \leq 3,73$. Mahasiswa yang memiliki $NMR MAT > 3,73$ umumnya berasal dari Program Studi non Matematika / Statistika.
2. Berdasarkan pohon regresi NMR S2 terdapat dua kelompok mahasiswa yang terbentuk dengan dugaan rata-rata $NMR S2 < 3,00$ atau DO yaitu :
 - 1). Kelompok mahasiswa yang berasal dari Perguruan Tinggi BHMN, $NMR S1 \leq 2,69$ serta pekerjaan bukan dosen maupun peneliti.
 - 2). Kelompok mahasiswa yang berasal dari Perguruan Tinggi Jawa Non BHMN dan Luar Jawa, $NMR STK \leq 2,61$.
3. Berdasarkan pohon regresi masa studi S2 terdapat dua kelompok mahasiswa yang terbentuk dengan dugaan rata-rata masa studinya ≤ 3 semester (Do) yaitu :
 - 1). Kelompok mahasiswa yang bukan sebagai dosen maupun peneliti, dan berasal dari program studi Matematika, Pendidikan Matematika dan Program studi lainnya.
 - 2). Kelompok mahasiswa yang bekerja sebagai dosen maupun peneliti, asal PT Jawa non BHMN dan Luar Jawa dan $NMR STK \leq 2,61$.
4. Pohon regresi juga menunjukkan bahwa mahasiswa yang tergolong berhasil dengan baik adalah kelompok mahasiswa yang bekerja

sebagai dosen maupun peneliti, berasal dari perguruan tinggi BHMN dan $NMR S1 > 2,85$.

5. Kelompok mahasiswa yang cenderung gagal (dropt out) adalah :
 - 1). Kelompok mahasiswa yang bukan sebagai dosen maupun peneliti.
 - 2). Kelompok mahasiswa yang bekerja sebagai dosen maupun peneliti, asal PT Jawa non BHMN dan Luar Jawa , $NMR STK \leq 2,61$.

Saran

Dari pohon regresi yang diperoleh disarankan untuk menyeleksi mahasiswa baru dengan memperhatikan peubah-peubah yang berpengaruh untuk menghasilkan mutu lulusan yang baik. Mahasiswa yang diterima di Program Studi Statistika SPs IPB yang diprediksikan berhasil menyelesaikan studi adalah mahasiswa yang pekerjaannya sebagai dosen atau peneliti, memiliki $NMR S1 \geq 2,69$ bagi mahasiswa yang berasal dari perguruan tinggi BHMN dan memiliki $NMR STK S1 \geq 2,61$ bagi mahasiswa yang berasal dari perguruan tinggi Jawa non BHMN dan Luar Jawa.

Penelitian ini mengkaji sisi input (masukan) dari suatu sistem pendidikan. Penelitian lebih lanjut dari sisi proses belajar mengajar diperlukan untuk mempelajari bagaimana keberhasilan mahasiswa jika dikaitkan dengan proses pembelajarannya. Apakah dengan input yang tetap, perbaikan proses pembelajaran dapat meningkatkan keberhasilan mahasiswa S2 Statistika IPB?

DAFTAR PUSTAKA

- Breiman, L. J. H. Friedman, R. A. Alshen, and C. J. Stone. 1993. Classification and Regression Trees. Chapman and Hall, New York
- Faridhan Y.E. 2003. Metode Klasifikasi Berstruktur Pohon Dengan Algoritma Cruise, Quest, dan Chaid [Tesis]. Bogor: Program Pascasarjana, Institut Pertanian Bogor.
- IPB. 2001. Katalog Program Pascasarjana IPB. Bogor: Institut Pertanian Bogor.
- Kudus, A. 1999. Penerapan Metode Regresi Berstruktur Pohon pada Pendugaan Masa Rawat Kelahiran Bayi (Studi Kasus di Rumah Sakit Hasan Sadikin Bandung) [Tesis]. Bogor: Program Pascasarjana, Institut Pertanian Bogor.

- Lewis, R. J. 2000. An Introduction to Classification and Regression Tree (CART) Analysis. Department of Emergency Medicine, Harbor-UCLA Medical Center, Torrance, California.
- Loh WY. 2002. Regression Trees with Unbiased Variable Selection and Interaction Detection, *Statistica Sinica* 12:361-386.
- Made S. 2002. Peubah-peubah Penciri yang Mempengaruhi Tingkat Keberhasilan Mahasiswa S2 STK IPB [Tesis]. Bogor: Program Pascasarjana, Institut Pertanian Bogor.
- Steinberg, Dan and Phillip Colla. 2000. CART for Windows User's Guide. Salford Systems, 2000.
- Thoha, I.F. 2003. Studi Tentang Tingkat Keberhasilan Mahasiswa S2 Program Pascasarjana IPB [Skripsi]. Bogor: Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor.

