

**PENERAPAN METODE KLASIFIKASI *RANDOM FOREST* DALAM  
MENGIDENTIFIKASI FAKTOR PENTING PENILAIAN MUTU PENDIDIKAN**

**APPLICATION OF *RANDOM FOREST* CLASSIFICATION METHOD IN IDENTIFYING  
IMPORTANT FACTORS OF EDUCATIONAL QUALITY**

**Aditya Ramadhan, Budi Susetyo, Indahwati**

**Institut Pertanian Bogor**

**Kampus IPB Dramaga, Bogor 16680**

**E-mail: email.bersama2@gmail.com, buset008@yahoo.com, indah.stk@gmail.com**

Naskah diterima tanggal: 12-05-2019 disetujui tanggal: 28-06-2019

**Abstract:** *National Education Standards serves as the basis of education development strategy based on the result of evaluation the implementation of education. The evaluation is implemented through accreditation and national exam. The objective of this study is to analyze the score of computer-based national exam based on accreditation scores per items of instrument by applying multiclass random forest classification modeling. The research used Computer-Based National Exam data in 2018 and accreditation data from the year of 2017 and 2018. This study employed random forest for multiclass classification. The results showed that, based on the evaluation model, classification accuracy value in multiclass random forest was 83.49%. In addition, this model produces important variables in classifying the average value of computer-based national examination, i.e., items laboratory conditions (x71, x68, x69, x67), electrical installation (x62), infrastructure (x64), canteen (x83), laboratory (x55), special service officers (x56), certified teachers (x39), library staff (x54), head of administration (x51), literacy activities for students (x33), use of textbooks (x14), and community/partner collaboration in education management (x96). Based on the indicators of important variables, National Education Standards that have important role are facility and infrastructure standards, educator and educational staff standards, and graduate competence standards. Therefore, improving the quality of education can be done by improving school facilities, the competency of teacher and education staff, and graduate competency.*

**Keywords:** *national education standards, national exam, multiclass random forest, variable importances*

**Abstrak:** *Standar Nasional Pendidikan (SNP) berfungsi sebagai dasar strategi pengembangan pendidikan berdasarkan hasil evaluasi pelaksanaan pendidikan. Evaluasi pelaksanaan pendidikan dilaksanakan melalui akreditasi dan ujian nasional (UN). Tujuan penelitian ini untuk menganalisis nilai ujian nasional berbasis komputer (UNBK) berdasarkan skor akreditasi per butir instrumen dengan menerapkan pemodelan klasifikasi random forest multikelas. Penelitian ini menggunakan data UNBK tahun 2018 dan data hasil akreditasi tahun 2017 dan 2018. Metode penelitian yang digunakan adalah pemodelan klasifikasi random forest multikelas. Hasil penelitian menunjukkan bahwa, pertama, berdasarkan evaluasi model, nilai akurasi klasifikasi dalam pemodelan klasifikasi random forest multikelas sebesar 83.49%. Kedua, model ini menghasilkan tingkat kepentingan variabel prediktor (butir-butir instrumen akreditasi) dalam mengklasifikasikan nilai rata-rata UNBK yakni kondisi laboratorium (x71, x68, x69, x67), instalansi listrik (x62), prasarana (x64), kantin (x83), kondisi laboran (x55), petugas layanan khusus (x56), guru tersertifikat (x39), tenaga perpustakaan (x54), kepala administrasi (x51), kegiatan literasi S/M bagi peserta didik*

(x33), penggunaan buku teks (x14), dan kerja sama masyarakat/mitra dalam pengelolaan pendidikan (x96). Berdasarkan indikator variabel penting tersebut, SNP yang memiliki peran penting adalah Standar Sarana dan Prasarana, Standar Pendidik dan Tenaga Kependidikan, dan Standar Kompetensi Lulusan. Oleh karena itu, peningkatan mutu pendidikan dapat dilakukan dengan meningkatkan sarana dan prasarana, kompetensi pendidik dan tenaga kependidikan, serta kompetensi lulusan.

**Kata kunci:** standar nasional pendidikan, ujian nasional, random forest multikelas, tingkat kepentingan variabel

## PENDAHULUAN

Upaya peningkatan mutu pendidikan secara nasional merupakan program pemerintah yang berkelanjutan dalam rangka menjawab berbagai tantangan di era globalisasi, termasuk kemajuan ilmu pengetahuan dan teknologi serta pergerakan tenaga ahli yang sangat masif. Mutu pendidikan merupakan tingkat kesesuaian antara penyelenggaraan pendidikan dengan Standar Nasional Pendidikan (SNP). SNP dikembangkan oleh Badan Standar Nasional Pendidikan (BSNP) yang merupakan kriteria minimal yang harus dipenuhi oleh satuan pendidikan dan/atau penyelenggara pendidikan dengan mempertimbangkan kondisi dan keragaman Indonesia. Peraturan Pemerintah (PP) Nomor 19 Tahun 2005, SNP terdiri dari Standar Kompetensi Lulusan (SKL), Standar Isi (SI), Standar Proses (SPR), Standar Pendidik dan Tenaga Kependidikan (SPT), Standar Sarana dan Prasarana (SSP), Standar Pengelolaan (SPL), Standar Pembiayaan (SB) dan Standar Penilaian (SPN) (Republik Indonesia, 2005).

Salah satu bentuk evaluasi penyelenggaraan pendidikan jenjang pendidikan dasar dan menengah berdasarkan SNP adalah penilaian yang dilaksanakan oleh Badan Akreditasi Nasional Sekolah/Madrasah (BAN S/M) dalam bentuk akreditasi. BAN S/M mengembangkan instrumen/perangkat akreditasi yang berisi butir-butir pernyataan hasil penilaian berdasarkan dokumen, pengamatan, dan verifikasi lapangan yang selanjutnya digunakan untuk memberikan penilaian terhadap delapan SNP. Instrumen untuk melakukan akreditasi jenjang SMA/MA

ditetapkan dalam Permendikbud Nomor 004/H/AK/2017 tentang Kriteria dan Perangkat Akreditasi SMA/MA.

SNP berfungsi sebagai dasar penyusunan strategi pengembangan mutu pendidikan berdasarkan hasil evaluasi belajar nasional seperti ujian nasional (UN). Menurut Permendikbud Nomor 3 Tahun 2017, UN adalah kegiatan pengukuran capaian kompetensi lulusan pada mata pelajaran tertentu secara nasional dengan mengacu pada SKL. SKL terdiri dari kriteria kualifikasi kemampuan peserta didik yang diharapkan dapat dicapai setelah menyelesaikan masa belajarnya. SKL digunakan sebagai acuan utama pengembangan SI, SPR, SPN, SPT, SSP, SPL, dan SB. Hasil UN dijadikan sebagai salah satu pertimbangan untuk pemetaan mutu program dan/atau satuan pendidikan, dasar seleksi masuk jenjang pendidikan berikutnya, serta pembinaan dan pemberian bantuan kepada satuan pendidikan dalam upaya meningkatkan mutu pendidikan (Kementerian Pendidikan dan Kebudayaan, 2017a).

Akreditasi sekolah/madrasah dan UN memiliki acuan yang sama yaitu SNP. Pelaksanaannya harus sejalan karena keduanya merupakan program penjaminan mutu berbasis standar. Hasil UN dan akreditasi dipergunakan sebagai acuan pada satuan pendidikan dalam upaya meningkatkan mutu layanan pendidikan (evaluasi pelaksanaan pendidikan). Dengan demikian, ketercapaian pemenuhan standar berdasarkan butir-butir instrumen akreditasi menjadi faktor penting di dalam penilaian mutu pendidikan.

Pemenuhan kelayakan layanan pendidikan berbasis standar setiap tahun senantiasa mengalami peningkatan dan usaha pemerataan selalu dikembangkan. Dukungan pemerintah pusat dan daerah selalu meningkat demi tercapainya pelayanan pendidikan yang adil dan merata. Merespon dukungan pemerintah dan kepedulian masyarakat dalam pelaksanaan pendidikan seyogyanya satuan pendidikan memanfaatkan sumber daya yang ada dengan melaksanakan pendidikan dengan baik dan berkualitas serta memanfaatkan hasil evaluasi pendidikan agar menghasilkan kompetensi lulusan yang berkualitas dan handal.

Penelitian tentang hubungan kausalitas antara SNP dan antara SNP dengan UN (khususnya UNBK) dilakukan oleh Ferezagia, Susetyo, Indriyanto, dan Pajajaran (2015) yang melihat hubungan kausalitas antara 8 SNP pada jenjang SMA dengan menggunakan metode *generalized structured component analysis* (GSCA) yang menyatakan bahwa SPN dan SPR berpengaruh signifikan terhadap SKL. Penelitian oleh Setiawan, Susetyo, & Fitrianto, (2018) dengan metode GSCA pada data akreditasi dan UNBK jenjang SMP/MTs yang menyimpulkan bahwa SKL, SPN, dan SPR berpengaruh signifikan terhadap UN. Hijrah, Susetyo, & Sartono, (2018) menyampaikan metode *Partial Least Square-Path Modeling* (PLS-PM) pada data SNP jenjang SMK yang menyimpulkan bahwa SB dan SPR berpengaruh signifikan terhadap SKL.

Berdasarkan penelitian Ferezagia, *et al.* 2015; Setiawan, *at al.*, 2018; dan Hijrah, *et al.*, 2018, bahwa UN dan akreditasi memiliki hubungan yang saling terkait dan memberikan pengaruh yang kuat. Hal tersebut akan mendorong untuk melakukan pengklasifikasian hubungan antarvariabel yang dapat menghasilkan tingkat kepentingan variabel dalam penilaian mutu pendidikan yakni dengan pemodelan klasifikasi *random forest*. Metode *random forest* merupakan pengembangan dari metode *Classification and Regression Tree*

(CART) dengan menerapkan metode *bootstrap aggregating* dan *random feature selection*. Kelebihan metode ini antara lain dapat menghasilkan akurasi yang lebih tinggi, dapat mengatasi data dalam jumlah yang besar secara efisien, dan tidak terdapat pemangkasan variabel seperti pada algoritma pohon klasifikasi tunggal (Breiman, 2001). Selain menghasilkan akurasi hasil prediksi yang sangat tinggi, *random forest* juga menghasilkan nilai tingkat kepentingan dari variabel prediktor dalam pengklasifikasian variabel respon (Sartono & Syafitri, 2010). Variabel prediktor yang digunakan sebagai pemilah dalam pengklasifikasian nilai rata-rata ujian nasional berbasis komputer (UNBK) akan memiliki nilai tingkat kepentingan tinggi dalam memengaruhi mutu pendidikan.

Permasalahan yang dirumuskan dalam penelitian ini: 1) Berapa banyak satuan pendidikan yang telah terakreditasi oleh BAN S/M yang memiliki kriteria mutu "Baik", "Cukup", dan "Kurang" berdasarkan nilai rata-rata UNBK; 2) Bagaimana menerapkan pemodelan klasifikasi *random forest* dalam mengidentifikasi faktor-faktor penting yang memengaruhi penilaian mutu pendidikan pada satuan pendidikan berdasarkan nilai UNBK dan hasil akreditasi. Tujuan dari penelitian ini yaitu untuk menganalisis nilai UNBK berdasarkan skor akreditasi per butir instrumen dengan menerapkan pemodelan klasifikasi *random forest* multikelas. Manfaat penelitian ini adalah memberikan informasi hasil evaluasi pelaksanaan pendidikan dalam upaya peningkatan mutu pendidikan dan perbaikan sistem penilaian akreditasi. Data analisis yang digunakan yaitu nilai rata-rata UNBK tahun 2018 dan hasil akreditasi jenjang SMA tahun 2017-2018.

## METODE

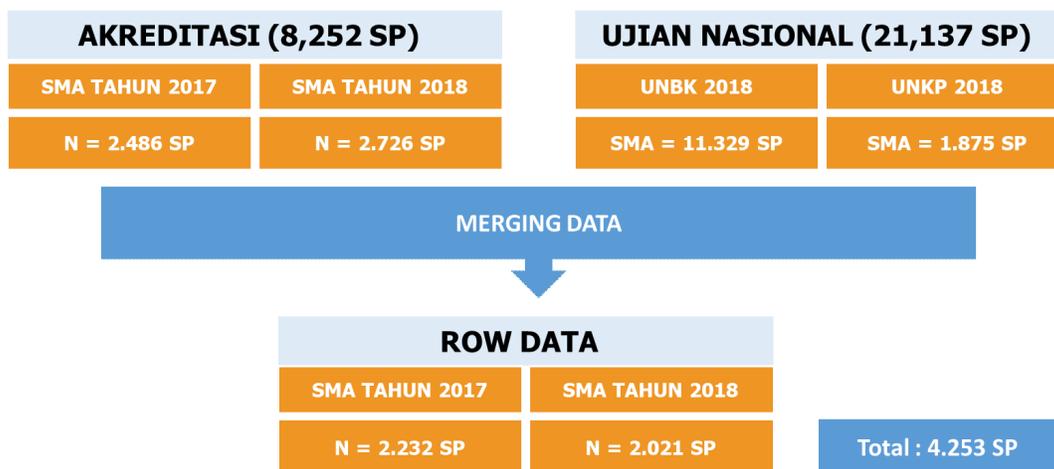
Penelitian ini menggunakan pendekatan kuantitatif dengan menggunakan pemodelan klasifikasi *random forest* untuk respon multi-

kategori (multikelas). Penelitian ini menggunakan data sekunder yang mengklasifikasikan nilai rata-ran UNBK berdasarkan hasil akreditasi berupa skor butir instrumen akreditasi. Unit penelitian adalah satuan pendidikan jenjang SMA seluruh Indonesia yang terakreditasi pada tahun 2017 dan 2018. Data yang digunakan dalam penelitian ini adalah data rata-ran hasil UNBK tahun 2018 untuk 3 mata uji utama Bahasa Indonesia (BIN), Bahasa Inggris (ING), dan Matematika (MTK), yang selanjutnya digunakan sebagai variabel respon. Data nilai hasil UNBK diperoleh dari Badan Penelitian dan Pengembangan (Balitbang) Kemendikbud. Nilai UNBK dikategorikan menjadi empat, yakni "Sangat Baik" ( $Y > 85$ ), "Baik" ( $70 < Y_d < 85$ ), "Cukup" ( $55 < Y_d < 70$ ), dan "Kurang" ( $Y_d < 55$ ). Untuk variabel prediktor adalah data hasil akreditasi tahun 2017-2018 dengan menggunakan 129 skor butir instrumen hasil akreditasi. Data akreditasi diperoleh dari BAN-S/M.

Data yang digunakan sebanyak 4.253 satuan pendidikan jenjang SMA. Data tersebut merupakan hasil penggabungan dari data hasil akreditasi SMA yang terdiri 5.212 satuan pendidikan dan data hasil UNBK yang terdiri 11.329 satuan pendidikan berdasarkan Nomor Pokok Sekolah Nasional (NPSN). Gambar 1 menunjukkan penggabungan data antara data hasil akreditasi tahun 2017 dan 2018 dengan data hasil UNBK tahun 2018 berdasarkan NPSN.

*Software* statistika yang digunakan dalam analisis adalah *software R*, yang merupakan *software open source*. Tahapan analisis yang dilakukan dalam data penelitian ini adalah sebagai berikut:

1. Melakukan eksplorasi data untuk memperoleh gambaran umum data.
2. Membagi data (4.253 satuan pendidikan) menjadi data latih (data yang digunakan untuk pembentukan model klasifikasi *random forest*) dan data uji (data yang digunakan sebagai evaluasi model yang dihasilkan) dengan berbagai opsi (7:3, 8:2, dan 9:1).
3. Melakukan penanganan data tidak seimbang akibat kategorisasi pada variabel respon yakni dengan melakukan pembobotan kelas.
4. Melakukan klasifikasi *random forest* multikelas pada data latih. Algoritma pemodelan klasifikasi *random forest* adalah pohon klasifikasi (*classification tree*) yang dapat digunakan untuk mengklasifikasi data dengan variabel respon multikelas (Zhou, Wang, & Fujita, 2016; Zhou, Wang, & Fujita, 2017): a) Tahap *bootstrap aggregating*, penarikan contoh acak dengan pengembalian berukuran  $n$  dari gugus data latih, yang kemudian disebut dataset, sehingga terdapat dataset sejumlah  $n$ . Pembentukan pohon klasifikasi dilakukan pada setiap dataset tanpa melakukan



Gambar 1 Gugus Data Penelitian (Data Akreditasi dan UNBK)

pemangkasan variabel prediktor; b) Tahap *random feature selection*, pemilahan variabel prediktor secara acak sebanyak  $m$ , dimana  $m < 129$  (umumnya dipilih  $1/3$  dari total variabel prediktor) di setiap pembentukan pohon klasifikasi; c) Langkah a dan b dilakukan sebanyak  $k$  kali sehingga didapatkan  $k$  buah pohon klasifikasi. Peluang untuk setiap kelas yang masuk ke dalam simpul akhir dari sebuah pohon klasifikasi adalah penduga/prediksi kelas di mana dengan nilai peluang terbesar;

5. Melakukan prediksi kelas/kategori variabel respon pada data uji. Prediksi kelas dari variabel respon suatu amatan adalah dengan menggunakan *majority voting* (nilai modus) dari hasil prediksi  $n$  pohon klasifikasi;
6. Mengevaluasi model klasifikasi dengan menghitung nilai akurasi. Nilai akurasi merupakan rasio dari data yang terklasifikasi benar dengan keseluruhan data uji.
7. Menghitung tingkat kepentingan variabel prediktor dalam pengklasifikasian nilai rataan UNBK berdasarkan butir-butir instrumen akreditasi dengan menghitung *Mean Decrease Gini* (MDG). MDG merupakan perhitungan rata-rata impuritas tereduksi dari setiap variabel prediktor ketika digunakan sebagai pemilah pada setiap pembentukan pohon tunggal klasifikasi pada *algoritma random forest* (Sandri & Zuccolotto, 2006).

## HASIL DAN PEMBAHASAN

### Eksplorasi Data

Pemenuhan SNP pada jenjang SMA pada tahun 2017 dan 2018 dapat dikategorikan baik dan sangat baik. Namun, hal ini tidak diikuti dengan

tingginya mutu pendidikan satuan pendidikan yang dicerminkan oleh nilai rataan UN. Pada Tabel 1 terlihat bahwa korelasi antara nilai UN untuk tiga matauji utama dengan pemenuhan SNP bernilai sangat kecil, bahkan untuk matauji matematika memiliki korelasi negatif terhadap enam SNP. Ini artinya kenaikan skor SNP memiliki pengaruh yang sangat kecil terhadap kenaikan pada nilai UN BIN dan ING, sedangkan untuk matauji MTK, penurunan skor SNP memiliki pengaruh yang sangat kecil terhadap kenaikan skor enam SNP. Hal ini tidak sesuai dengan teori pendidikan yang menyebutkan bahwa akreditasi dan UN merupakan bentuk evaluasi pelaksanaan mutu yang berbasis standar, dimana setiap kenaikan skor SNP seharusnya dapat menghasilkan kenaikan nilai UN (Badan Standar Nasional Pendidikan, 2018). Oleh karena itu, pada penelitian ini digunakan data UNBK yang lebih baik dalam mencerminkan mutu pendidikan dimana skor SNP memiliki korelasi positif terhadap nilai UNBK dan memiliki nilai korelasi yang cukup besar (Tabel 2).

Data yang digunakan pada penelitian ini adalah 4.253 satuan pendidikan jenjang SMA yang terdiri dari 1.894 SMAN dan 2.359 SMA Swasta. Nilai korelasi antara skor SNP dengan nilai UNBK yang cukup besar dan bernilai positif memiliki arti bahwa kenaikan skor SNP akan memengaruhi kenaikan pada nilai UNBK. Tabel 2 menunjukkan bahwa nilai korelasi antara mata uji (BIN, ING, dan MTK) memiliki nilai korelasi yang sangat tinggi, nilai korelasi yang terendah terdapat pada korelasi antara MTK dengan BIN sebesar 0.73. Nilai korelasi antardelapan SNP secara umum memiliki nilai korelasi yang cukup tinggi dengan nilai korelasi terbesar yakni korelasi antara SI dengan SPR sebesar 0.86.

Tabel 1 Matriks Korelasi Skor SNP dengan nilai hasil UNKP

	SI	SPR	SKL	SPT	SSP	SPL	SB	SPN
BIN	0.16	0.15	0.14	0.12	0.16	0.15	0.14	0.14
ING	0.05	0.05	0.06	0.04	0.06	0.06	0.07	0.05
MTK	0.00	-0.02	-0.03	-0.03	-0.05	-0.01	0.01	-0.01

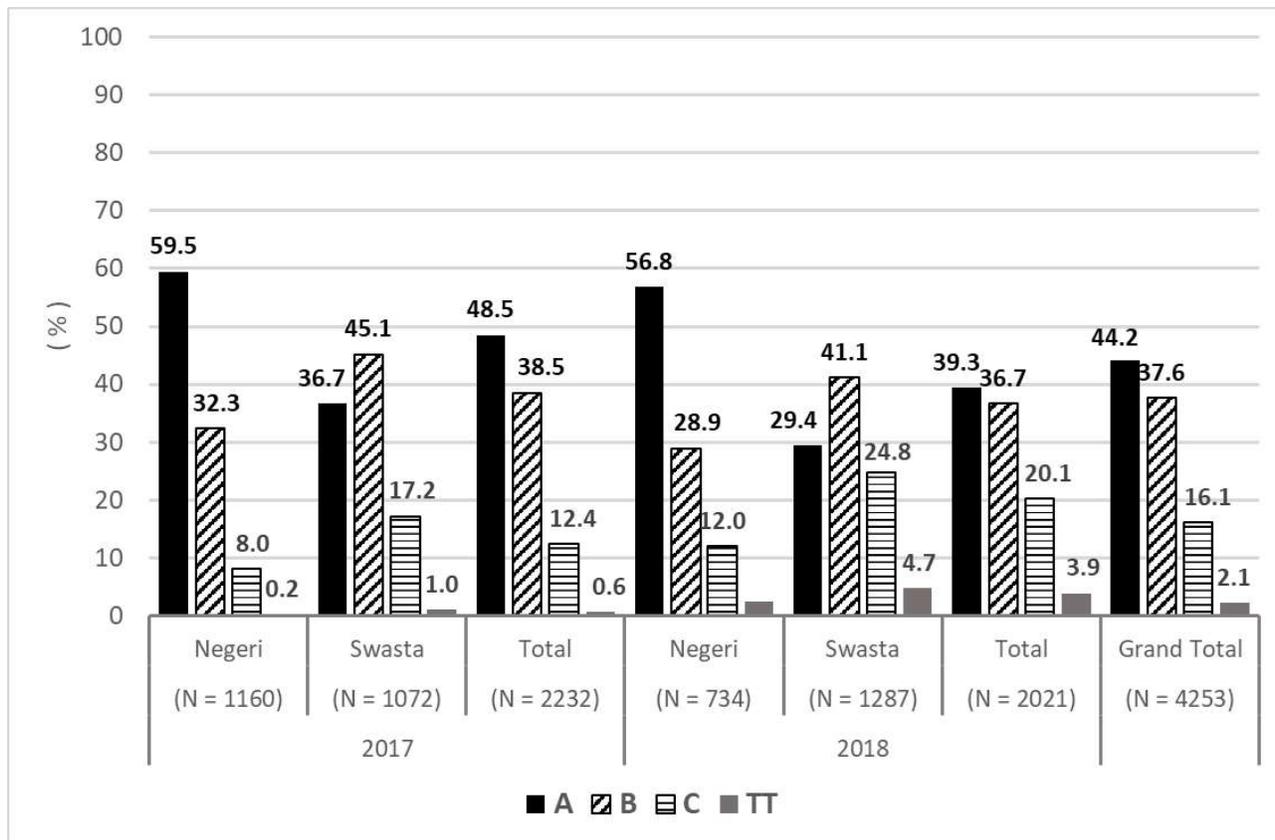
Tabel 2 Matriks Korelasi Skor SNP dengan Nilai Hasil UNBK

	BIN	ING	MTK	SI	SPR	SKL	SPT	SSP	SPL	SB	SPN
BIN	1	0.811	0.677	0.379	0.394	0.426	0.412	0.511	0.395	0.316	0.365
ING	0.811	1	0.808	0.340	0.375	0.400	0.401	0.495	0.376	0.277	0.349
MTK	0.677	0.808	1	0.226	0.255	0.275	0.301	0.364	0.257	0.180	0.232
SI	0.379	0.340	0.226	1	0.858	0.803	0.655	0.673	0.802	0.707	0.817
SPR	0.394	0.375	0.255	0.858	1	0.847	0.737	0.737	0.826	0.693	0.826
SKL	0.426	0.400	0.275	0.803	0.847	1	0.724	0.739	0.813	0.680	0.801
SPT	0.412	0.401	0.301	0.655	0.737	0.724	1	0.805	0.728	0.595	0.651
SSP	0.511	0.495	0.364	0.673	0.737	0.739	0.805	1	0.751	0.617	0.676
SPL	0.395	0.376	0.257	0.802	0.826	0.813	0.728	0.751	1	0.738	0.817
SB	0.316	0.277	0.180	0.707	0.693	0.680	0.595	0.617	0.738	1	0.703
SPN	0.365	0.349	0.232	0.817	0.826	0.801	0.651	0.676	0.817	0.703	1

Sedangkan, nilai korelasi antara delapan SNP dan hasil UNBK menunjukkan nilai yang cukup beragam. Namun keseluruhan nilai korelasi kurang dari 0.5. Korelasi tertinggi terdapat pada SSP dengan keempat mata uji dengan nilai korelasi lebih dari 0.41. Hal ini menunjukkan bahwa SSP memiliki pengaruh yang tinggi

dibandingkan SNP lainnya jika melihat nilai korelasi.

Gambar 2 menunjukkan persentase status akreditasi berdasarkan jenis dan status satuan pendidikan. Secara keseluruhan, jumlah satuan pendidikan yang terakreditasi A sebesar 44,2%, terakreditasi B 37,6%, terakreditasi C 16,1%,



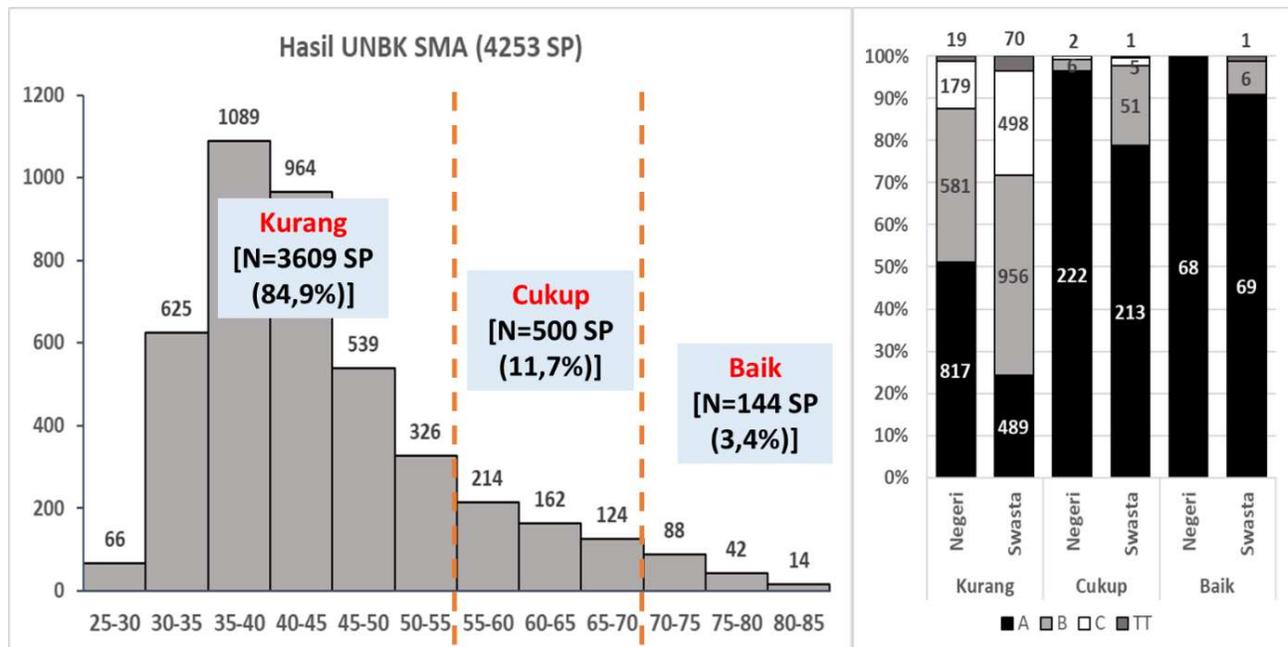
Gambar 2 Status Akreditasi berdasarkan Jenis dan Status Satuan Pendidikan

dan Tidak Terakreditasi (TT) 2,1%. Berdasarkan status satuan pendidikan, SMAN pada tahun 2017 dan 2018 cenderung mendapatkan peringkat akreditasi A dengan persentase 59,5% dari 1.160 satuan pendidikan dan 56,8% dari 734 satuan pendidikan. Sedangkan SMA swasta pada tahun 2017 dan 2018 cenderung memperoleh peringkat akreditasi B dengan persentase 45,1% dari 1.072 satuan pendidikan dan 41,1% dari 1.287 satuan pendidikan. Hal ini dapat disimpulkan bahwa sebagian besar sekolah SMA yang menerapkan UNBK merupakan sekolah dengan akreditasi A dan B (81.8%).

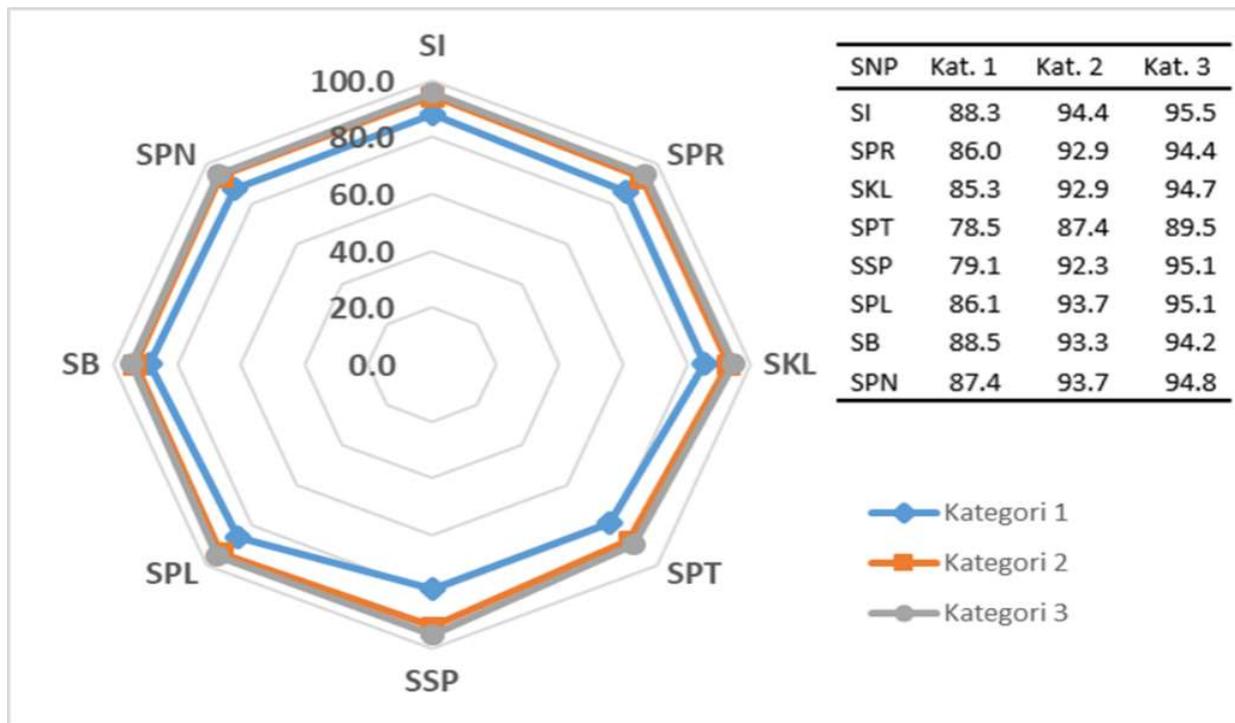
Kategorisasi mutu pendidikan berdasarkan nilai UNBK dibagi menjadi tiga kategori karena untuk kategori 4 memiliki amatan yang sangat sedikit sehingga kategori 4 diklasifikasikan ke dalam kategori 3. Gambar 3 merupakan grafik kategorisasi mutu pendidikan, dimana satuan pendidikan dengan mutu pendidikan yang "kurang baik" sebanyak 3609 satuan pendidikan (84,9%), mutu pendidikan yang "cukup baik" 500 satuan pendidikan (11,7%), dan 144 satuan pendidikan (3,4%) bermutu "baik". Kategori mutu pendidikan "kurang" masih didominasi satuan pendidikan yang terakreditasi A dan B

baik untuk satuan pendidikan berstatus negeri maupun swasta (SMAN kategori A: 51% dan B: 37%; SMAS kategori A: 24% dan B: 46%). Hal tersebut menunjukkan bahwa nilai rata-rata UNBK per satuan pendidikan jenjang SMA cenderung mengalami hasil yang kurang baik. Hasil yang kurang baik tersebut ada kaitannya dengan integritas peserta didik (pelaksanaan UNBK) serta tingkat kesulitan soal-soal UNBK yang memasukkan beberapa soal dengan daya nalar tinggi atau *higher order thinking skill* 10-15% dari total soal (Pusat Penilaian Pendidikan, 2018).

Capaian pemenuhan delapan SNP yang meningkat di tingkat satuan pendidikan cenderung akan meningkatkan level kategorisasi mutu pendidikan berdasarkan nilai rata-rata UNBK. Hal tersebut dapat terlihat pada Gambar 4 dimana satuan pendidikan dengan pencapaian delapan SNP yang tinggi cenderung memiliki tingkatan level kategorisasi UNBK yang tinggi (nilai rata-rata UNBK yang baik). Hal ini dapat disimpulkan bahwa hasil akreditasi secara bersama-sama berpengaruh terhadap pencapaian nilai UNBK. Hal tersebut selaras dengan penelitian Raharjo (2014) yang menunjukkan



Gambar 3 Hasil Akreditasi Berdasarkan Kategorisasi Mutu Pendidikan (Nilai UNBK)



Gambar 4 Kategorisasi UNBK berdasarkan Capaian Delapan SNP

bahwa delapan SNP secara bersama-sama berpengaruh terhadap pencapaian nilai UN.

### Penanganan Data Tidak Seimbang

Data tidak seimbang merupakan suatu kejadian dimana terdapat kelas data yang memiliki amatan jauh lebih sedikit dibanding kelas lainnya (Bunkhumpornpat, Sinapiromsaran, & Lursinsap, 2012). Dalam algoritma pemodelan klasifikasi yang digunakan cenderung tidak memerhatikan ketidakseimbangan data sehingga kurang memadai jika terdapat kasus data tidak seimbang. Akibat dari kejadian ini, kelas minoritas akan mengalami kesalahan klasifikasi. Penanganan terhadap data tidak seimbang ini perlu diminimalisir kesalahan klasifikasi.

Pada penelitian ini, kelas mayoritas adalah yang memiliki nilai UNBK berkategori "kurang" dengan persentase 84.9%, sedangkan untuk 2 kategori lainnya termasuk kelas minoritas yakni kategori "cukup baik" dengan 11,7%, dan kategori "baik" hanya 3,4% dari 6.771 satuan pendidikan. Salah satu cara menangani data tidak seimbang adalah dengan melakukan

pembobotan kelas. Pada amatan di kelas mayoritas diberikan bobot yang lebih kecil dibandingkan dengan amatan pada kelas minoritas (kelas minoritas diberi bobot 1) sehingga memberi efek pengamatan pada kelas minoritas meningkat dan menjadi seimbang dengan kelas mayoritas. Pembobot kelas merupakan dasar lanjutan dalam pembobotan parameter model klasifikasi yang dihasilkan.

### Penerapan *Random Forest* Multikelas

Penelitian ini menggunakan *random forest* multikelas dengan *package* "mlr" (*Machine Learning in R*) dan untuk algoritma *random forest* menggunakan "classif.ranger" yang dapat bekerja lebih cepat untuk pengimplementasian klasifikasi *random forest* untuk data berdimensi tinggi (Toomey, 2014). Pada algoritma *random forest* multikelas dataset dibagi menjadi data latih dan data uji. Data latih digunakan untuk pembentukan model klasifikasi, sedangkan data uji sebagai bahan mengevaluasi model. *Cut off point* dalam membagi gugus data ditentukan sebesar 90:10, 80:20, dan 70:30. Penentuan

beberapa *cut off* dilakukan untuk melihat seberapa sensitif kinerja kedua metode dalam mengklasifikasikan mutu pendidikan berdasarkan persentase pembagian dataset sehingga diperoleh nilai evaluasi yang optimal.

Pembentukan pohon klasifikasi dilakukan dengan melakukan pemilahan secara berulang dan membagi sekumpulan data menjadi dua sekatan. Pemilahan ini dilakukan dengan memilih pemilah terbaik yang berasal dari semua kemungkinan pemilahan. Variabel prediktor pada penelitian ini merupakan variabel kategori berskala ordinal, maka terdapat *v-1* kemungkinan pemilahan. Pemilihan pemilah terbaik dari masing-masing variabel prediktor dilakukan dengan menggunakan nilai impuritas. Nilai impuritas mengukur keheterogenan suatu simpul. Pemilahan yang memaksimalkan nilai penurunan keheterogenan di dalam masing-masing simpul anak relatif terhadap simpul utama merupakan pemilah terbaik. Indeks *gini* merupakan nilai yang digunakan untuk mendefinisikan ukuran dari fungsi impuritas. Peluang untuk setiap kelas yang masuk ke dalam simpul akhir dari sebuah pohon klasifikasi adalah penduga/prediksi kelas dimana dengan nilai peluang terbesar (Hüllermeier & Vanderlooy, 2010)

Evaluasi model digunakan untuk mengukur kinerja pemodelan klasifikasi dengan matriks konfusi. Matriks konfusi merupakan tabel klasifikasi yang diperoleh dari jumlah ketepatan hasil prediksi dengan data aktual pada masing-masing amatan pada data uji (Suresha, Danti, & Narasimhamurthy, 2014). Ukuran kinerja model klasifikasi dalam mengevaluasi model menggunakan nilai akurasi klasifikasi (Kohavi, 1995). Akurasi merupakan rasio jumlah amatan yang diklasifikasikan dengan benar (prediksi sesuai dengan aktual) terhadap jumlah ke-seluruhan amatan yang diujikan (Adnan & Islam, 2015).

Hasil evaluasi model *random forest* multikelas pada Tabel 3 menunjukkan bahwa terdapat akurasi klasifikasi terbesar pada model klasifikasi *random forest* multikelas DF dengan

*cut off* data latih dan data uji 90:10. Nilai akurasi klasifikasi *random forest* multikelas *cut off* 90:10 sebesar 83.49%. Interval kesalahan dalam menghitung nilai akurasi model dengan batas bawah 79.61% dan batas atas nilai akurasi model klasifikasi sebesar 86.90%.

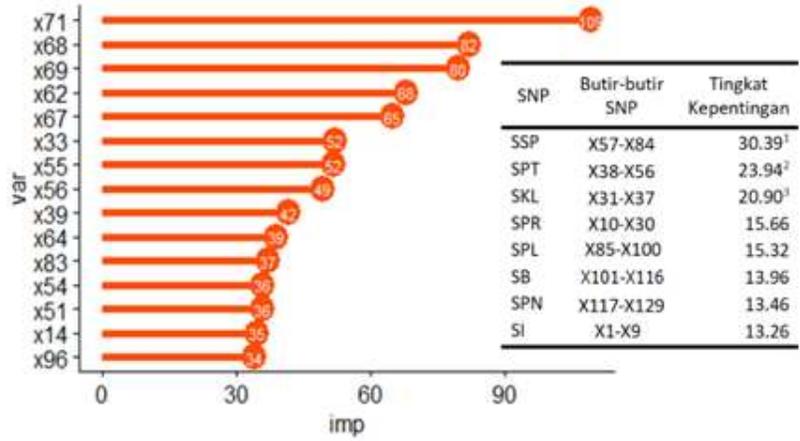
Tabel 3 Hasil Evaluasi Model Random Forest Multikelas

Evaluasi	<i>Cut Off</i>		
	90:10	80:20	70:30
Akurasi	83.49	83.04	82.78
BB	79.61	80.34	79.69
BA	86.90	85.50	86.71

Keterangan: BB: Batas Bawah; BA: Batas Atas

### Tingkat Kepentingan Variabel Prediktor

Pemodelan *random forest* menghasilkan tingkat kepentingan dari variabel prediktor (faktor penting) yang didapatkan dengan menghitung MDG. Gambar 5 merupakan grafik yang menjelaskan tingkat kepentingan variabel prediktor (ditampilkan 15 variabel dengan nilai MDG tertinggi). Variabel-variabel yang memiliki tingkat kepentingan tertinggi antara lain butir instrumen mengenai kondisi laboratorium bahasa (x71), laboratorium fisika (x68), laboratorium kimia (x69), instalansi listrik (x62), laboratorium biologi (x67), kegiatan literasi S/M dalam pembentukan sikap pembelajar sejati sepanjang hayat bagi peserta didik (x33), ketersediaan laboran sesuai ketentuan kualifikasinya (x55), ketersediaan petugas layanan khusus sesuai ketentuan kualifikasinya (x56), jumlah guru yang tersertifikat (x39), ketersediaan prasarana yang lengkap dan baik (x64), ketersediaan kantin sesuai ketentuan (x83), kualifikasi tenaga perpustakaan sesuai ketentuan (x54), ketersediaan kepala administrasi sesuai ketentuan kualifikasinya (x51), penggunaan buku teks dalam pembelajaran (x14), dan kerja sama masyarakat/mitra dalam pengelolaan pendidikan (x96).

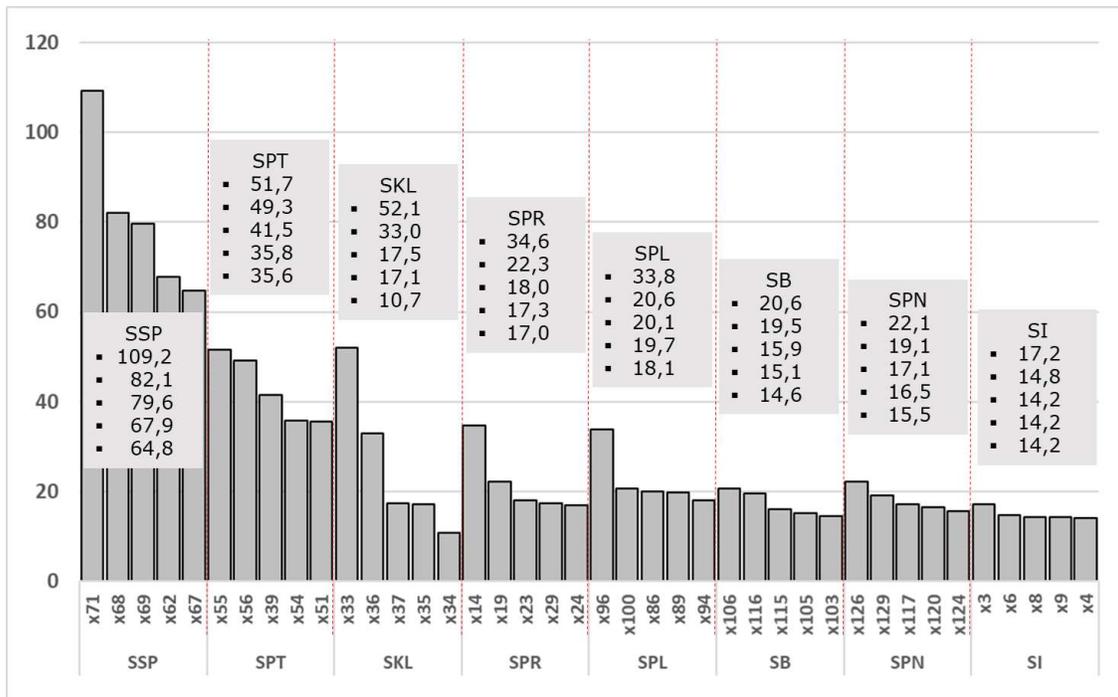


Gambar 5 Tingkat Kepentingan Variabel pada *Random Forest*

Tingkat kepentingan variabel dapat menggambarkan tingkat kepentingan delapan SNP dengan menghitung rata-rata tingkat kepentingan butir-butir instrumen dari penilaian delapan SNP. Tingkat kepentingan SNP dalam pengklasifikasian mutu pendidikan dapat terlihat pada Gambar 5. Gambar 5 menunjukkan bahwa SSP memiliki tingkat kepentingan tertinggi dalam pengklasifikasian mutu pendidikan di tingkat satuan pendidikan secara berurutan SPT, SKL, SPR, SPL, SB, SPN, dan SI. Hal ini senada dengan hasil penelitian yang dilakukan oleh Raharjo (2018) yang menunjukkan bahwa SPT, SSP, dan

SKL secara parsial memiliki pengaruh secara signifikan terhadap mutu satuan pendidikan (rata-rata UNBK).

Gambar 6 menunjukkan lima variabel prediktor dengan tingkat kepentingan tertinggi di setiap SNP yang menjadi faktor penting dalam mengklasifikasikan mutu pendidikan. Faktor penting yang menjadi pemilah dalam mengklasifikasikan mutu pendidikan (nilai UNBK) untuk SSP adalah ketersediaan dan kesesuaian laboratorium bahasa (x71), laboratorium fisika (x68), laboratorium kimia (x69), instalansi listrik (x62), dan laboratorium biologi (x67). Lima faktor



Gambar 6 Lima Variabel Prediktor dengan Tingkat Kepentingan Tertinggi per SNP

terpenting untuk SPT adalah ketersediaan laboran yang memenuhi ketentuan (x55), indikator ketersediaan petugas layanan khusus (x56), guru bersertifikat (x39), kualifikasi tenaga perpustakaan (x54), dan Kepala Administrasi sesuai ketentuan (x51). Untuk SKL, lima faktor terpenting adalah kegiatan literasi S/M dalam pembentukan sikap pembelajar sejati sepanjang hayat bagi peserta didik (x33), keterampilan berpikir peserta didik (x36), keterampilan bertindak peserta didik (x37), pengetahuan peserta didik (x35), dan kegiatan kesiswaan S/M dalam membentuk sikap sehat jasmani dan rohani peserta didik (x34).

Hasil analisis penerapan metode klasifikasi *random forest* dalam mengidentifikasi faktor penting dalam penilaian mutu pendidikan berdasarkan butir-butir instrumen akreditasi memberikan manfaat bagi pemerintah dan satuan pendidikan. Hal tersebut dilakukan dengan memberikan informasi butir-butir instrumen akreditasi yang penting untuk diperbaiki dalam upaya peningkatan mutu pendidikan dan landasan evaluasi dalam memperbaiki sistem penilaian akreditasi. Menurut Hendarman (2013), jika akreditasi sekolah tidak diimplementasikan untuk meningkatkan mutu pendidikan, maka sumber daya besar yang dikeluarkan dalam proses akreditasi tak sebanding dengan yang didapatkan. Hal tersebut juga disebutkan oleh Handayani (2016) bahwa hasil akreditasi dapat dijadikan umpan balik dalam usaha pemberdayaan dan pengembangan kinerja warga satuan pendidikan dalam rangka menerapkan visi, misi tujuan, sasaran, strategi dan program satuan pendidikan. Di samping itu, hasil akreditasi diharapkan mampu memotivasi satuan pendidikan untuk terus meningkatkan mutu pendidikan secara bertahap, terencana, dan kompetitif baik di tingkat daerah, nasional, bahkan di tingkat internasional.

Kontribusi pemenuhan delapan SNP dengan melihat tingkat kepentingan variabel maupun

tingkat kepentingan SNP dalam peningkatan mutu pendidikan. Hal tersebut dapat diartikan bahwa satuan pendidikan secara nasional memiliki ketercapaian pemenuhan SSP, SPT, dan SKL yang sangat beragam atau memiliki ketimpangan yang cukup besar dalam pengategorisasian mutu pendidikan (rata-rata UNBK). Sedangkan, standar yang lainnya cenderung tidak jauh berbeda di setiap satuan pendidikan.

Program prioritas baik yang telah, sedang, maupun yang akan dilakukan oleh pemerintah yang berkaitan dengan pemerataan layanan pendidikan melalui pemenuhan SSP yaitu: program bantuan operasional sekolah (BOS), pembangunan sarpras (sekolah baru, ruang kelas baru, laboratorium), rehab gedung, dan bantuan buku teks. Program terkait pemerataan pemenuhan dan peningkatan SPT antara lain: sertifikasi guru, guru 3T (Terdepan, Terluar, dan Tertinggal), peningkatan kompetensi guru (melalui beasiswa, bimtek, pelatihan), dan portal rumah belajar. Program terkait SKL, di antaranya: program revolusi mental dengan pendidikan karakter, implementasi kurikulum 2013, portal rumah belajar, dan gerakan literasi nasional.

Sistem penilaian akreditasi, baik kebijakan pelaksanaan akreditasi maupun pengembangan/penyempurnaan instrumen akreditasi, setiap tahunnya dievaluasi oleh BAN-S/M. Pada tahun 2019, kebijakan akreditasi mengalami perubahan orientasi dari penilaian dokumentasi dan lapangan menjadi penilaian akreditasi berbasis performa. Penilaian akreditasi berbasis performa ini merupakan penilaian akreditasi yang bersifat dokumentasi. Dan tidak lagi menjadi input penilaian akreditasi yang dilakukan satuan pendidikan. Data dokumentasi satuan pendidikan terinput otomatis sesuai data yang terintegrasi dengan data pokok pendidikan (Dapodik), data penjaminan mutu pendidikan (PMP), dan data *education management information system* (EMIS) Kementerian Agama. Penilaian akreditasi berbasis performa juga akan dilakukan dengan

menilai tingkat kepuasan pemangku kepentingan dan pengakuan masyarakat terhadap satuan pendidikan.

## SIMPULAN DAN SARAN

### Simpulan

Berdasarkan pembahasan dapat disimpulkan bahwa kategori mutu pendidikan "kurang" masih didominasi satuan pendidikan yang terakreditasi A dan B baik untuk satuan pendidikan berstatus negeri maupun swasta. Hal tersebut menunjukkan bahwa nilai rata-rata UNBK per satuan pendidikan jenjang SMA cenderung mengalami hasil yang kurang baik. Hasil yang kurang baik tersebut ada kaitannya dengan integritas peserta didik terkait pelaksanaan UNBK. Tingkat kesulitan soal-soal UNBK yang memasukkan beberapa soal dengan daya nalar tinggi atau *higher order thinking skill* 10-15% dari total soal juga ada kaitannya dengan hasil yang masih kurang.

Pemodelan klasifikasi *random forest* multikelas dalam mengklasifikasikan mutu pendidikan berdasarkan hasil UNBK dan skor akreditasi per butir instrumen pada jenjang SMA tahun 2018 memberikan tingkat akurasi model secara keseluruhan sebesar 83.49%. Berdasarkan tingkat kepentingan per butir instrumen dalam pengklasifikasian mutu pendidikan yang dihasilkan pada pemodelan klasifikasi *random forest*, Standar Nasional Pendidikan yang memiliki peran penting dalam pengklasifikasian mutu pendidikan yaitu Standar Sarana dan Prasarana, Standar Pendidikan dan Tenaga Kependidikan, dan Standar Kompetensi Lulusan. Sementara Standar Proses, Standar Pengelolaan, Standar Biaya, Standar Penilaian, dan Standar Isi memiliki peran dengan tingkat kepentingan yang lebih kecil dalam mengidentifikasi mutu pendidikan. Hal ini dapat diartikan bahwa hasil akreditasi merupakan prediktor dalam mengklasifikasikan mutu pendidikan, namun masih belum sepenuhnya dapat menjadi prediktor dalam mengidentifikasi

mutu pendidikan yang tepat di tingkat satuan pendidikan.

### Saran

Berdasarkan simpulan yang dihasilkan dapat disarankan kepada pemerintah baik pusat maupun daerah, dan satuan pendidikan bersama-sama saling berkolaborasi untuk berusaha meningkatkan: 1) pemenuhan Standar Sarana Prasarana yaitu terkait ketersediaan dan kesesuaian laboratorium bahasa, laboratorium fisika, laboratorium kimia, laboratorium biologi, dan instalansi listrik; 2) pemenuhan Standar Pendidik dan Tenaga Kependidikan yaitu terkait ketersediaan dan kesesuaian dari tenaga laboran, petugas layanan khusus, tenaga perpustakaan, kepala administrasi serta ketersediaan guru yang bersertifikat; dan 3) pemenuhan Standar Kompetensi Lulusan yaitu kegiatan literasi S/M dalam pembentukan sikap pembelajar sejati sepanjang hayat dan pengembangan keterampilan berpikir bagi peserta didik. Peningkatan pemenuhan ketiga Standar Nasional Pendidikan ini dilakukan secara terencana dan berkesinambungan. Bagi sekolah yang belum memenuhi delapan standar, evaluasi pelaksanaan pendidikan tidak hanya memperhatikan pemenuhan ketiga Standar Nasional Pendidikan tersebut, tetapi juga memperhatikan pemenuhan delapan Standar Nasional Pendidikan dalam meningkatkan pencapaian mutu pendidikan.

Penelitian selanjutnya disarankan untuk dapat mengembangkan model klasifikasi dengan menambahkan variabel prediktor yang diindikasikan memengaruhi pengklasifikasian mutu pendidikan, seperti data pokok pendidikan dasar dan menengah (dapodikdasmen) dan/atau menggunakan data penjaminan mutu pendidikan (PMP). Perlu dikembangkan *binarization random forest* dan pemodelan klasifikasi variabel ganda dalam mengidentifikasi faktor penting dalam mengklasifikasikan mutu pendidikan.

## PUSTAKA ACUAN

- Adnan, M.N. & Islam, M.Z. (2015). One-vs-all binarization technique in the context of random forest. *Journal of Computational Intelligence Machine Learning*, 5(1), 22–24.
- Badan Standar Nasional Pendidikan. (2018). Kebijakan pelaksanaan ujian nasional tahun 2018. Jakarta: Badan Penelitian dan Pengembangan, Kementerian Pendidikan dan Kebudayaan.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Bunkhumpornpat, C., Sinapiromsaran, K., & Lursinsap, C. (2012). DBSMOTE: density-based synthetic minority over-sampling technique. *International Journal of Research on Intelligent Systems for Real Life Complex Problems (Applied Intelligence)*, 36(3), 664–684. <https://doi.org/10.1007/s10489-011-0287-y>.
- Ferezagia, D.V., Susetyo, B., Indriyanto, B., & Pajajaran, J. (2015). Generalized Structured Component Analysis (GSCA) for National Education Standards (NES) of secondary school in Indonesia. *Global Journal of Pure and Applied Mathematics*, 11(4), 2441–2449.
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Prosiding of the 14th International Joint Conference on Artificial Intelligence - 2*, 1137–1143.
- Handayani, M. (2016). Pencapaian standar nasional pendidikan berdasarkan hasil akreditasi SMA di Provinsi DKI Jakarta. *Jurnal Pendidikan dan Kebudayaan*, 1(2), 179-201.
- Hendarman. (2013). Pemanfaatan hasil akreditasi dan kredibilitas asesor sekolah/madrasah. *Jurnal Pendidikan dan Kebudayaan*, 19(4), 532-542.
- Hijrah, M., Susetyo, B., & Sartono, B. (2018). Structural equation modeling of national standard education of vocational high school using partial least square path modeling. *International Journal of Scientific Research in Science, Engineering and Technology*, 4(4), 1418–1422.
- Hüllermeier, E. & Vanderlooy, S. (2010). Combining predictions in pairwise classification/ : an optimal adaptive voting strategy and its relation to weighted voting. *Pattern Recognit*, 43(1), 128–142.
- Pusat Penilaian Pendidikan. (2018). Hasil Ujian Nasional Tahun 2018. Jakarta: Badan Penelitian dan Pengembangan, Kementerian Pendidikan dan Kebudayaan.
- Raharjo, S.B. (2014). Kontribusi delapan standar nasional pendidikan terhadap pencapaian prestasi belajar. *Jurnal Pendidikan dan Kebudayaan*, 20(4), 470-482.
- Raharjo, S.B. (2018). Capaian standar nasional pendidikan sebagai prediktor mutu sekolah. *Jurnal Pendidikan dan Kebudayaan*, 3(2), 129-140.
- Republik Indonesia. (2005). Peraturan Pemerintah Nomor 19 tahun 2005 tentang Standar Nasional Pendidikan. Jakarta: Kementerian Pendidikan dan Kebudayaan.
- Kementerian Pendidikan dan Kebudayaan. (2017). Peraturan Menteri Pendidikan dan Kebudayaan Nomor 3 Tahun 2017 tentang Penilaian Hasil Belajar oleh Pemerintah dan Penilaian Hasil Belajar oleh Satuan Pendidikan.
- Sartono, B. & Syafitri, U.D. (2010). Metode pohon gabungan: Solusi pilihan untuk mengatasi kelemahan pohon regresi dan klasifikasi tunggal. *Forum Statistika dan Komputasi*, 15(1), 1–7.

- Sandri, M. & Zuccolotto, P. (2006). Variable selection using random forest. Dalam S. Zani, A. Cerioli, M. Riani, & M. Vichi (Eds), *The Classification and Data Analysis Group (CLADAG) of the Italian Statistical Society: Data Analysis, Classification and the Forward Search*. New York (US): Springer Berlin Heidelberg.
- Setiawan, I.A., Susetyo, B., & Fitrianto, A. (2018). Application of generalized structural component analysis to identify relation between accreditation and national assessment. *International Journal of Scientific Research in Science, Engineering and Technology*, 4(10), 93–97.
- Suresha, M., Danti, A. & Narasimhamurthy, S.K. (2014). Decision trees to multiclass prediction for analysis of arecanut data. *Journal of Computer Systems Science and Engineering*, 29(1), 105-114.
- Toomey, D. (2014). *R for data science*. Brimingham: Packt Publishing Ltd.
- Zhou, L., Wang, Q. & Fujita, H. (2016). Predicting the listing status of chinese listed companies with multi-class classification models. *International Journal of Informatics and Computer Science Intelligent Systems Applications*, 328, 222–236.
- Zhou, L., Wang, Q., & Fujita, H. (2017). One versus one multi-class classification fusion using optimizing decision directed acyclic graph for predicting listing status of companies. *International Journal on Multi-Sensor, Multi-Source Information Fusion*, 36(1), 80–89.