

# SYSTEMATIC LITERATURE REVIEW PREDIKSI KINERJA SISWA: TREN PENELITIAN, METODE, DATASET, DAN ATRIBUT

Avira Budianita<sup>a</sup>, M. Adhitya Wardhana<sup>b</sup>, Fandy Indra Pratama<sup>c</sup>

[avirabudianita@umkudus.ac.id](mailto:avirabudianita@umkudus.ac.id), [muhammadadhityawardhana@gmail.com](mailto:muhammadadhityawardhana@gmail.com), [fandy@unwahas.ac.id](mailto:fandy@unwahas.ac.id)

<sup>a</sup>Sistem Informasi, Universitas Muhammadiyah Kudus  
Purwosari, Kudus, Indonesia

<sup>b</sup>Bisnis Digital, Universitas Muhammadiyah Kudus  
Purwosari, Kudus, Indonesia

<sup>d</sup>Teknik Informatika, Universitas Wahid Hasyim  
Menoreh, Semarang, Indonesia

---

## Abstrak

Prediksi kinerja siswa banyak diteliti oleh peneliti-peneliti dunia khususnya di bidang pendidikan dan data mining. Prediksi kinerja siswa pada beberapa penelitian pasti tidak lepas dari metode data mining khususnya metode klasifikasi. Maraknya penelitian tentang kinerja siswa karena dataset yang digunakan yaitu dataset pendidikan yang mana pasti memiliki kapasitas yang besar dan terdapat beberapa yang belum terolah dengan baik. Dari penelitian prediksi kinerja siswa, terdapat beberapa permasalahan yaitu belum teridentifikasinya secara baik tren penelitian, metode, dataset, dan atribut dalam prediksi kelulusan siswa. Tren penelitian tentang prediksi kinerja siswa yang ada masih belum cukup untuk mengetahui tingkat popularitas dari penelitian prediksi kinerja siswa dan masih lemahnya identifikasi mengenai metode, dataset dan atribut yang sering digunakan dan terbaik. Kurangnya analisis tersebut menjadikan systematics literature review ini jawaban untuk mengetahui tren penelitian prediksi kinerja siswa serta metode, dataset, dan atribut apa saja yang digunakan dalam prediksi kinerja siswa. Dari systematics literature review ini didapatkan hasil analisis berupa tren penelitian yaitu tren tahun penelitian dan tren negara penelitian serta hasil analisis berupa metode, dataset, dan atribut yang sering digunakan dan terbaik dalam prediksi kinerja siswa.

**Kata Kunci:** Systematic Literature Review, Kinerja Siswa, Data Mining, Dataset Pendidikan, Tren Penelitian

## Abstract

*Student performance predictions are widely researched by world researchers, especially in the fields of education and data mining. Prediction of student performance in several studies certainly can not be separated from data mining methods, especially classification methods. The rise of research on student performance because the dataset used is an educational dataset which must have a large capacity and there are some that have not been processed properly. From research predictions of student performance, there are several problems that are not well identified research trends, methods, datasets, and attributes in predicting student graduation. Research trends on student performance predictions are still not enough to know the level of popularity of student performance prediction research and the weak identification of methods, datasets and attributes that are often used and the best. This lack of analysis makes systematics literature review the answer to find out trends in student performance prediction research and what methods, datasets, and attributes are used in predicting student performance. From the systematics literature review, it is obtained the results of analysis in the form of research trends, namely research trends and research country trends, and analysis results in the form of methods, datasets, and attributes that are often used and the best in predicting student performance.*

**Keywords:** Systematic Literature Review, Student Performance, Data Mining, Educational Dataset, Research Trends

---

## I. PENDAHULUAN

Memprediksi kinerja siswa menjadi penelitian yang populer pada saat ini yang dilakukan oleh para peneliti di ranah data mining. Hal tersebut dikarenakan data

pendidikan yang sangat besar dan belum terolah secara maksimal. Dengan banyaknya data tersebut memungkinkan para pengurus akademik untuk menjadi bahan penentuan khususnya dalam penentuan kinerja siswa

yang mana kinerja siswa dalam sebuah instansi harus selalu dimonitor agar tetap menjaga kualitas lulusan dan instansi itu sendiri.

Kinerja siswa dapat didasarkan dari beberapa faktor-faktor yang beragam seperti variabel nilai seperti nilai pelajaran atau nilai mata kuliah; variabel data pribadi siswa seperti jenis kelamin, asal kota, asal sekolah, dan lain-lain; variabel sosial-ekonomi seperti gaji orang tua, pendidikan orang tua, lingkungan rumah, dan sebagainya; variabel psikologis seperti memiliki gadget, transportasi ke sekolah, dan lain sebagainya (Ramesh, Parkavi and Ramar, 2013). Dari beberapa data tersebut memungkinkan pihak sekolah atau perguruan tinggi untuk memonitor kinerja siswanya. Beberapa faktor juga dapat berpengaruh untuk menentukan pencapaian yang akan diraih oleh siswa di masa mendatang (Jananto, 2013).

Data mining menjadi teknik yang populer dalam analisis prediksi kinerja siswa (Shahiri, Husain and Rashid, 2015). Teknik data mining digunakan untuk membangun model berdasarkan data yang belum diketahui akan digunakan untuk mengidentifikasi informasi baru (Osmanbegović and Suljić, 2012). Metode yang ada pada data mining yang beragam dapat membantu peneliti dalam mengolah data yang sangat banyak dan hasil dari pengolahan tersebut yang berupa informasi baru dapat dijadikan bahan analisis. Pada prediksi kinerja siswa, para peneliti sebagian besar menggunakan metode data mining klasifikasi. Klasifikasi adalah teknik data mining prediksi, membuat prediksi tentang nilai-nilai data menggunakan hasil yang diketahui dari data yang berbeda (Bhardwaj, 2011). Metode klasifikasi dan prediksi pada data mining memiliki beberapa algoritma seperti Decision Tree yang di dalamnya terdapat ID3, C4.5, J48, dan CART serta Bayesian yang di dalamnya terdapat Naive Bayes dan Bayesian Network (Yadav and Pal, 2012).

Banyak metode, dataset, dan atribut berbeda yang digunakan dalam penelitian dan sangat kompleks, dengan demikian Figurean menyeluruh tentang keadaan penelitian prediksi kinerja siswa saat ini

belum terlihat. Review sistematis ini bertujuan untuk mengidentifikasi dan menganalisa tren penelitian, metode, dataset, dan atribut yang digunakan dalam prediksi kinerja siswa pada penelitian antara 2010 hingga 2016.

Dalam makalah ini, metodologi penelitian dijelaskan pada bab 2. Hasil dan jawaban dari pertanyaan penelitian disajikan pada bab 3. Terakhir hasil dari penelitian ini diringkas pada bab terakhir.

## II. METODE PENELITIAN

### A. Review Metode

Pendekatan sistematis dipilih untuk meninjau literatur tentang prediksi kinerja siswa. Systematic Literature Review (SLR) didefinisikan sebagai proses mengidentifikasi, menilai dan menafsirkan semua bukti penelitian yang tersedia dengan tujuan untuk menyediakan jawaban untuk pertanyaan penelitian secara spesifik (Kitchenham *et al.*, 2009).

### B. Research Questions

Research Questions (RQ) atau Pertanyaan penelitian ditetapkan untuk menjaga review agar tetap terfokus. RQ dirancang dengan bantuan kriteria: Population, Intervention, Comparison, Outcomes, and Context (PICOC) (Budgen and Brereton, 2006).

**Table 1.** Ringkasan PICOC

Population	Kinerja siswa
Intervention	Prediksi kinerja siswa, prediksi, klasifikasi, metode, atribut, dataset
Comparison	Metode, atribut, dataset prediksi kinerja siswa
Outcomes	Keakuratan prediksi kinerja siswa, metode prediksi kinerja siswa terbaik
Context	Lembaga-lembaga akademik. Semua jenis studi empirik.

Pertanyaan penelitian dan motivasi yang akan dibahas pada kajian pustaka ini akan disajikan dalam Table 2.

**Table 2.** Research Questions pada Literature Review

ID	Research Question	Motivation
RQ1	Siapa peneliti yang paling aktif	Mengidentifikasi para peneliti yang paling aktif dan

	berpengaruh di bidang prediksi kinerja siswa?	berpengaruh yang telah berkontribusi banyak pada area penelitian prediksi kinerja
RQ2	Pada tahun penelitian prediksi kinerja siswa menjadi tren studi?	Mengidentifikasi tahun dimana prediksi kinerja siswa sering dipelajari oleh para peneliti
RQ3	Di negara mana saja penelitian prediksi kinerja siswa menjadi tren studi?	Mengenali negara mana saja prediksi kinerja siswa sering dipelajari oleh para peneliti
RQ4	Metode apa yang paling sering digunakan untuk prediksi kinerja siswa?	Mengidentifikasi metode yang paling banyak digunakan untuk prediksi kinerja siswa
RQ5	Metode yang terbaik ketika digunakan dalam prediksi kinerja siswa?	Mengidentifikasi perbaikan metode yang diusulkan dalam prediksi kinerja siswa
RQ6	Jenis atribut apa saja yang digunakan untuk prediksi kinerja siswa?	Mengidentifikasi perbaikan atribut yang diusulkan untuk prediksi kinerja siswa
RQ7	Jenis atribut apa saja yang paling sering digunakan untuk prediksi kinerja siswa?	Mengidentifikasi atribut yang paling banyak digunakan untuk prediksi kinerja siswa
RQ8	Jenis dataset tingkat pendidikan apa saja yang paling sering digunakan untuk prediksi kinerja siswa?	Mengidentifikasi dataset tingkat pendidikan yang paling banyak digunakan untuk prediksi kinerja siswa

Studi utama pada review ini, metode, atribut, dan dataset prediksi kinerja siswa digunakan untuk jawaban RQ4 sampai RQ9 yang diekstrak dari pertanyaan-pertanyaan yang spesifik diantara RQ4 hingga RQ9. Sedangkan RQ1 hingga RQ3 akan membantu memberikan Figurean trend penelitian tentang prediksi kinerja siswa.

Figure 1 menunjukkan pemikiran dasar dari systematic literature review. Tujuan utama dari systematic literature review ini adalah untuk mengidentifikasi metode, atribut dan dataset yang digunakan dalam prediksi kinerja siswa.

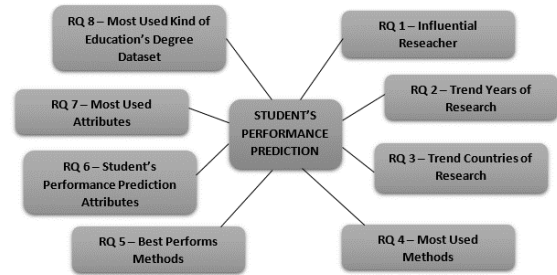


Figure 1. Peta Pemikiran Dasar dari SLR Prediksi Kinerja Siswa

### C. Strategi Pencarian

Proses pencarian dimulai dari pemilihan perpustakaan digital. Sebelum memulai pencarian, seperangkat database harus dipilih untuk meningkatkan kemungkinan menemukan artikel yang sangat relevan. Berikut adalah daftar database digital pencarian:

- ACM Digital Library ([dl.acm.org](http://dl.acm.org))
- IEEE Explore ([ieeexplore.ieee.org](http://ieeexplore.ieee.org))
- ScienceDirect ([sciencedirect.com](http://sciencedirect.com))
- Springer ([springerlink.com](http://springerlink.com))

Database yang dicari merupakan judul, kata kunci dan abstrak. Pencarian dibatasi pada tahun publikasi: 2010-2016. Dua jenis publikasi yang dimasukkan dalam review ini yaitu jurnal dan prosiding konferensi.

### D. Pemilihan Studi

Paket perangkat lunak Mendeley (<http://mendeley.com>) digunakan untuk menyimpan dan mengelola hasil pencarian pemilihan artikel. Daftar akhir dari pemilihan artikel tersebut yaitu sebanyak 35 artikel. Kemudian keseluruhan teks dari 35 artikel tersebut dianalisis. Tujuan melakukan analisis untuk mengetahui tingkat relevansi artikel tersebut dengan review yang akan dilakukan.

### E. Ekstraksi Data

Artikel yang telah dipilih pada langkah sebelumnya yang diambil untuk mengumpulkan data yang berkontribusi untuk menjawab pertanyaan-pertanyaan penelitian yang bersangkutan dalam review

ini. Ekstraksi data dirancang untuk mengumpulkan data dari artikel yang diperlukan untuk menjawab pertanyaan penelitian. Enam properti yang digunakan untuk menjawab pertanyaan-pertanyaan penelitian yang ditunjukkan dalam tabel 4.

**Table 3.** Properti Ekstraksi Data yang dipetakan ke Research Questions

Property	Research Questions
Peneliti atau Author	RQ1
Tren Penelitian	RQ2, RQ3
Metode Prediksi Kinerja Siswa	RQ4, RQ5, RQ6
Atribut Prediksi Kinerja Siswa	RQ7, RQ8
Dataset Prediksi Kinerja Siswa	RQ9

#### F. Studi Penilaian Kualitas dan Sintesis data

Penilaian kualitas studi dapat digunakan untuk memandu interpretasi dari temuan sintesis dan untuk menentukan kekuatan kesimpulan yang diuraikan. Tujuan dari sintesis data adalah keseluruhan bukti dari artikel yang dipilih untuk menjawab pertanyaan-pertanyaan penelitian. Data yang diambil dalam tinjauan ini termasuk data kuantitatif dan data kualitatif. Beberapa alat visualisasi, termasuk grafik batang, diagram lingkaran, tabel, dan juga digunakan untuk meningkatkan presentasi dari distribusi metode dan data akurasi dari prediksi kinerja siswa.

#### G. Ancaman terhadap Validitas

Review ini bertujuan untuk menganalisis studi tentang prediksi kinerja siswa yang didasarkan pada statistik dan machine learning. Pencarian tidak didasarkan pada membaca judul semua karya secara manual yang diterbitkan dalam jurnal. Ini berarti bahwa ulasan ini mungkin telah dikeluarkan di beberapa dokumen prediksi kinerja siswa dari beberapa prosiding konferensi ataupun jurnal.

Pada review ini membahas tentang trend penelitian, metode, atribut, dan dataset secara mendetail. Berbeda dengan review sejenis yang hanya mereview metode dan atribut saja seperti pada review yang dilakukan oleh

Shahiri et.al (Shahiri, Husain and Rashid, 2015).

### III. HASIL DAN PEMBAHASAN

#### A. Peneliti yang Paling Aktif dan Berpengaruh

Dari proses pemilihan studi, dapatkan 35 jurnal dan prosiding terkait prediksi kinerja siswa yang berkontribusi dalam analisis review untuk menjawab pertanyaan penelitian yang sudah dibuat sebelumnya.

Dari jurnal dan prosiding yang dipilih, peneliti yang berkontribusi sangat baik dan yang sangat aktif di bidang prediksi kinerja siswa selanjutnya diidentifikasi. Figure 2 menunjukkan peneliti paling aktif dan berpengaruh di bidang prediksi kinerja siswa.

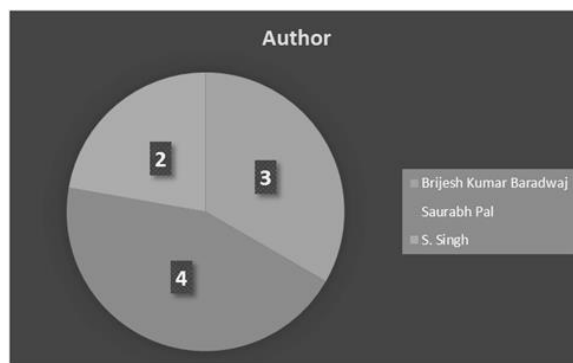


Figure 2. Peneliti yang Paling Aktif dan Berpengaruh

Pada Figure 2, didapati bahwa peneliti Brijesh Kumar Baradwaj yang berasal dari India melakukan penelitian tentang prediksi kinerja siswa terbanyak yaitu 4 penelitian dibanding peneliti lain. Selain itu Saurabh Pal menduduki peringkat dua dengan dengan 3 penelitian disusul dengan S. Singh dengan 2 penelitian. Selain ketiga peneliti tersebut hanya melakukan satu kali penelitian tentang prediksi kinerja siswa ataupun menjadi non-first author.

#### B. Tren Tahun Penelitian

Sebuah gambaran dari studi distribusi terkait prediksi kinerja siswa selama tahun ke tahun ditunjukkan pada Figure 3. Pada review ini tahun penelitian yang diambil adalah mulai dari tahun 2010 hingga terbaru yaitu 2016.

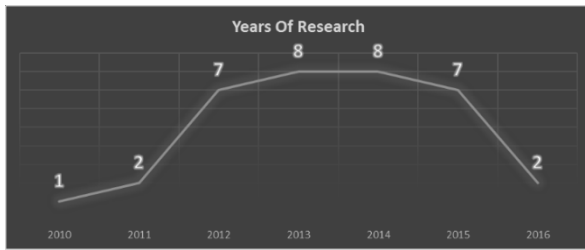


Figure 3. Tren Tahun Penelitian

Dari Figure 3, menunjukkan bahwa bidang penelitian tentang prediksi kinerja siswa mengalami kenaikan dalam kurun waktu antara tahun 2010 hingga 2013 meskipun mengalami penurunan pada tahun 2014 hingga 2016 dan penelitian terkait prediksi kinerja siswa masih menjadi penelitian yang relevan saat ini.

### C. Tren Negara Penelitian

Sama seperti tren tahun penelitian, tren negara penelitian juga menjadi pembahasan pada review ini. Dari tren negara penelitian ini dapat diketahui negara mana saja yang paling banyak dan berkontribusi terkait penelitian tentang prediksi kinerja siswa.

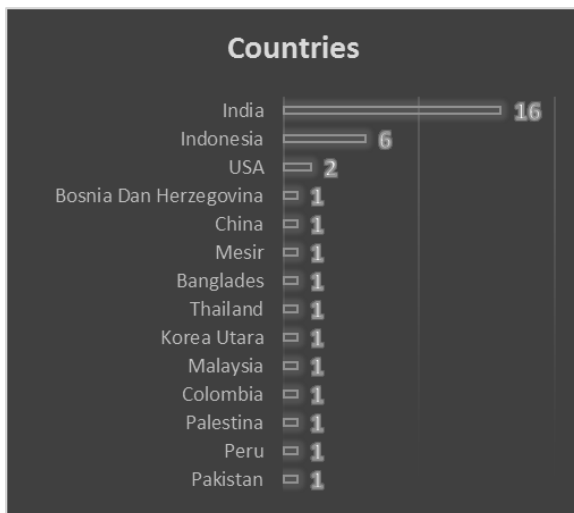


Figure 4. Tren Negara Penelitian

Figure 4 memperlihatkan tren negara penelitian prediksi kinerja siswa paling terbanyak berada pada negara India sebanyak 16 penelitian dalam kurun waktu 2010 hingga 2016. Kemudian disusul oleh Indonesia dengan 6 penelitian dan USA dengan 2 penelitian terkait prediksi kinerja siswa.

### D. Metode yang Paling Sering Digunakan dalam Prediksi Kinerja Siswa

Meninggalkan tentang tren penelitian tentang prediksi kinerja siswa, Pada poin ini

membahas tentang metode yang digunakan dalam prediksi kinerja siswa. Metode-metode ini merupakan metode dari teknik data mining.

Dari sekian banyak teknik data mining yang ada saat ini, Naive Bayes menempati urutan pertama algoritma atau metode yang paling sering digunakan dalam prediksi kinerja siswa meskipun pada Teorema Bayes juga terdapat Bayesian Network tetapi metode tersebut tidak terlalu sering digunakan dalam prediksi kinerja siswa. Metode Naive Bayes dikenal sebagai metode yang sederhana dalam hal klasifikasi dan prediksi (Ridwan, Suyono and Sarosa, 2013). Menyusul Naive Bayes terdapat Metode ID3, C4.5 dan Decision Tree yang mana ketiga algoritma tersebut masih dalam satu rumpun Decision Tree. Algoritma Decision Tree juga sering digunakan dalam prediksi kinerja siswa. Decision Tree juga merupakan metode yang populer dalam pengolahan data di bidang pendidikan (Quadri and Kalyankar, 2010).

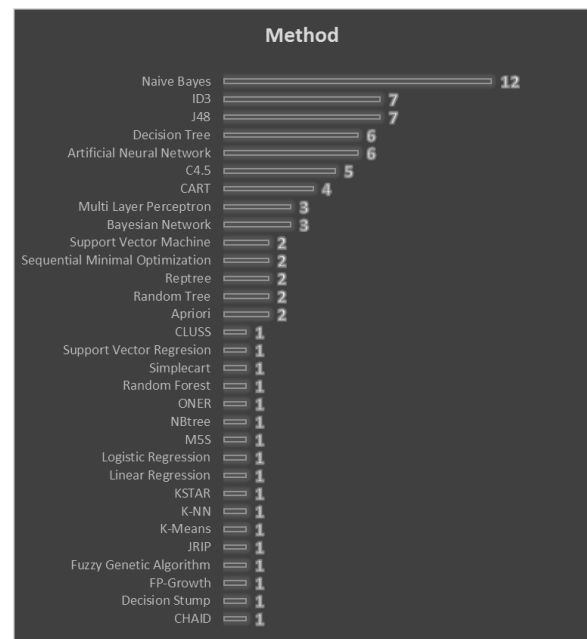


Figure 5. Metode yang Paling Sering Digunakan dalam Prediksi Kinerja Siswa

Pada Figure 5 dapat dilihat diagram metode yang sering digunakan dalam prediksi kinerja siswa.

### E. Metode Terbaik yang Digunakan dalam Prediksi Kinerja Siswa

Pada poin sebelumnya telah dijelaskan metode yang paling sering digunakan pada

prediksi kinerja siswa, tetapi hal tersebut tidak lantas membuat metode tersebut juga merupakan metode yang baik dalam hal tingkat akurasi maupun tingkat kesalahan minimum. Tingkat akurasi dan kesalahan minimum dapat dihitung manual maupun menggunakan tools data mining yaitu RapidMiner dan WEKA.

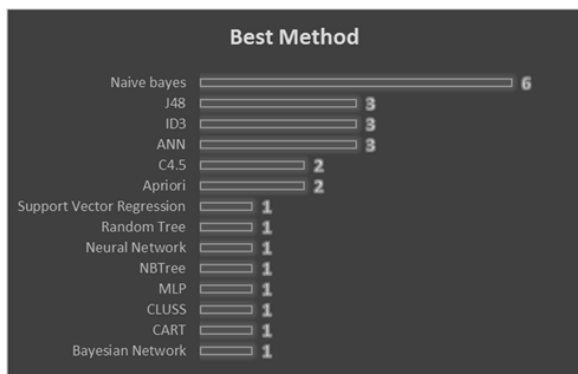


Figure 6. Metode Terbaik yang Digunakan dalam Prediksi Kinerja Siswa

Pada Figure 6 diagram menunjukkan bahwa Naive Bayes selain menjadi metode yang paling sering digunakan tetapi juga menjadi metode yang paling baik dalam hal akurasi pada prediksi kinerja siswa. Begitu pula dengan J48 dan ID3 yang tidak hanya menjadi metode yang paling sering digunakan tetapi juga menjadi metode yang terbaik pada prediksi kinerja siswa.

#### F. Atribut yang Digunakan dalam Prediksi Kinerja Siswa

Tidak hanya metode, atribut juga menjadi pembahasan penting dalam review prediksi kinerja siswa. Pada review serupa juga membahas tentang atribut apa saja yang digunakan dalam proses prediksi kinerja siswa (Shahiri, Husain and Rashid, 2015). Terdapat tiga faktor atribut yang mempengaruhi prediksi kinerja siswa yaitu Nilai, Data Personal Siswa, dan Kondisi Sosial-Ekonomi siswa (Ramesh, Parkavi and Ramar, 2013). Ketiga faktor atribut tersebut menjadi faktor terkuat dalam proses prediksi kinerja siswa.

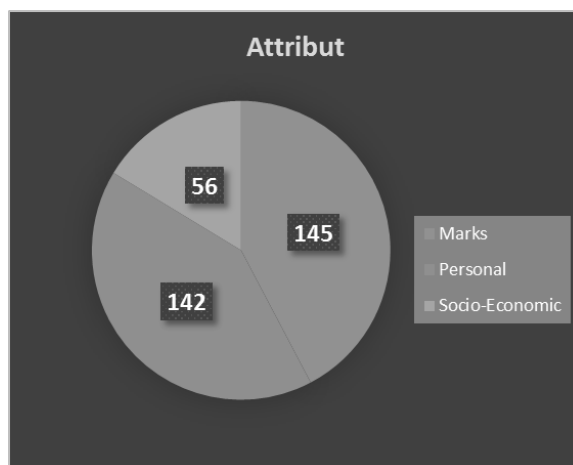


Figure 7. Atribut yang Digunakan dalam Prediksi Kinerja Siswa

Dari ketiga faktor atribut utama yang mempengaruhi prediksi kinerja siswa didapati bahwa faktor atribut nilai menjadi faktor atribut yang paling sering digunakan dalam prediksi kinerja siswa. Kemudian paling sering digunakan kedua yaitu data personal siswa dan disusul kondisi sosial-ekonomi siswa yang sering menjadi atribut pendukung pada prediksi kinerja siswa.

#### G. Atribut yang Paling Sering Digunakan dalam Prediksi Kinerja Siswa

Dari ketiga faktor atribut yang disampaikan pada poin 3.6, terdapat atribut-atribut yang paling sering digunakan dari masing-masing faktor atribut tersebut. Faktor atribut nilai berisi nilai-nilai siswa dari nilai mata pelajaran maupun mata kuliah, nilai rata-rata siswa dan nilai-nilai lainnya. Sedangkan faktor atribut data personal siswa berisi data-data diri siswa seperti jenis kelamin, asal kota, asal sekolah, dan data-data diri lainnya. Dan faktor atribut kondisi sosial-ekonomi siswa berisi data-data tentang kondisi sosial dan ekonomi siswa seperti jumlah saudara, jarak yang ditempuh ke sekolah atau kuliah, kepemilikan barang berharga, pendidikan dan gaji orang tua, dan data-data sosial-ekonomi siswa lainnya.

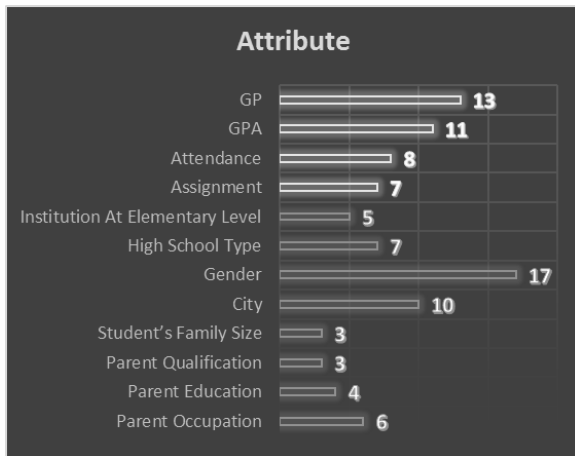


Figure 8. Atribut yang Paling Sering Digunakan dalam Prediksi Kinerja Siswa

Figure 8 menyajikan atribut yang sering digunakan dari ketiga faktor atribut nilai, data personal siswa, dan kondisi sosial-ekonomi siswa.

Dari faktor atribut nilai didapati empat peringkat teratas atribut yang sering digunakan yaitu GP atau indeks prestasi, GPA atau indeks prestasi kumulatif, attendance atau jumlah kehadiran siswa, assignment atau nilai tugas siswa. GP atau indeks prestasi menjadi atribut yang sering digunakan dari faktor atribut nilai dalam prediksi kinerja siswa.

Selanjutnya dari faktor atribut data personal siswa didapati empat peringkat teratas atribut yang sering digunakan yaitu institution at elementary level atau asal lembaga sekolah, high school type atau tipe asal sekolah, gender atau jenis kelamin, dan city atau kota asal. Gender atau jenis kelamin menjadi atribut yang sering digunakan dari faktor atribut data personal siswa dalam prediksi kinerja siswa.

Terakhir dari faktor atribut kondisi sosial-ekonomi siswa didapati empat peringkat teratas atribut yang sering digunakan yaitu student's family size atau ukuran keluarga siswa, parent qualification atau kualifikasi orang tua siswa, parent education atau pendidikan terakhir orang tua siswa, dan parent occupation atau pekerjaan orang tua siswa. Parent occupation atau pekerjaan orang tua siswa menjadi atribut yang sering digunakan dari faktor atribut kondisi sosial-ekonomi siswa dalam prediksi kinerja siswa

## H. Jenis Dataset Tingkat Pendidikan yang Digunakan dalam Prediksi Kinerja Siswa

Penelitian tentang prediksi kinerja siswa pasti membutuhkan dataset pendidikan terutama data siswa. Dari dataset tersebut nantinya akan diketahui apakah data-data tersebut dapat berkontribusi secara baik dalam prediksi kinerja siswa. Dari studi literatur yang sudah dilakukan, dataset yang digunakan semuanya menggunakan dataset privat yang didapatkan langsung di sekolah maupun universitas.

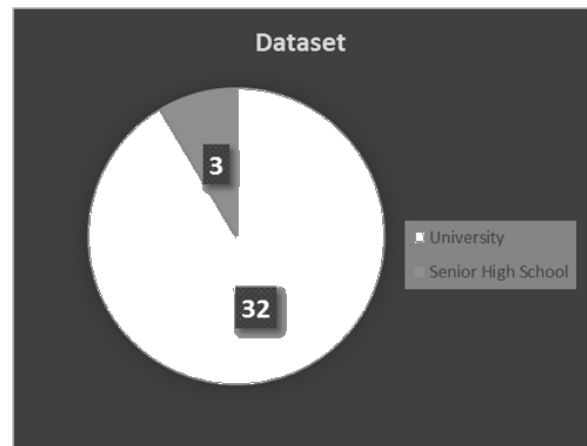


Figure 9. Jenis Dataset Tingkat Pendidikan yang Digunakan dalam Prediksi Kinerja Siswa

Dataset yang banyak digunakan oleh peneliti di bidang prediksi kinerja siswa merupakan dataset siswa di sekolah menengah atas maupun dataset mahasiswa di universitas. Dari Figure 9, terlihat bahwa dataset yang berasal dari universitas menjadi dataset yang paling banyak digunakan dalam penelitian prediksi kinerja siswa dibanding dataset siswa dari sekolah menengah atas. Beberapa peneliti menggunakan dataset yang berasal dari universitas karena disamping data yang banyak dan kompleks juga terdapat banyak sekali celah dalam pengolahan data mahasiswa. Data mahasiswa yang terolah dengan baik juga dapat menentukan capaian mahasiswa di masa yang akan mendatang.

## IV. KESIMPULAN DAN FUTURE WORK

Literature Review ini bertujuan untuk mengidentifikasi dan menganalisis tren penelitian, metode, atribut, dan dataset tentang prediksi kinerja siswa dari tahun 2010 hingga 2016. Dari hasil pencarian studi

didapati 35 jurnal dan proceeding berkaitan dengan prediksi kinerja siswa yang dipublikasi dari tahun 2010 hingga 2016. Kemudian Literature Review ini dikemas menjadi Systematics Literature Review. Systematics Literature Review didefinisikan sebagai proses identifikasi, menilai, dan menafsirkan semua bukti penelitian yang tersedia dengan tujuan untuk memberikan jawaban untuk pertanyaan penelitian tertentu.

Penelitian pada review ini dimulai dari identifikasi tren penelitian dengan menganalisis peneliti yang aktif dan berkontribusi dalam prediksi kinerja siswa, tahun tren penelitian dan negara tren penelitian. Dari identifikasi tersebut didapati bahwa peneliti yang aktif dan berkontribusi dalam prediksi kinerja siswa yaitu Brijesh Kumar Baradwaj yang berasal dari India. Kemudian tren tahun penelitian berada pada tahun 2010 hingga 2014 yang mengalami kenaikan meskipun pada tahun 2014-2016 mengalami penurunan tetapi penelitian tentang prediksi kinerja siswa masih menjadi penelitian yang relevan pada saat ini.

Identifikasi metode juga menjadi fokus identifikasi pada review tentang prediksi kinerja siswa ini. Dari hasil identifikasi didapati bahwa Naive Bayes yang disebut sebagai metode klasifikasi paling sederhana menjadi metode atau algoritma yang paling sering digunakan dan menjadi metode dengan akurasi terbaik dalam penelitian kinerja siswa.

Selain metode, identifikasi juga difokuskan pada atribut yang digunakan dalam prediksi kinerja siswa. Atribut yang sering digunakan dikemas dalam tiga faktor atribut yaitu faktor atribut nilai, data personal siswa, dan kondisi sosial-ekonomi siswa. Dari hasil identifikasi dan analisis didapatkan bahwa faktor atribut nilai menjadi faktor atribut yang paling sering digunakan dalam prediksi kinerja siswa. Dari ketiga faktor atribut tadi juga terdapat atribut-atribut yang paling sering digunakan pada prediksi kinerja siswa. Pada faktor atribut nilai, GP atau indeks prestasi menjadi atribut yang paling sering digunakan pada prediksi kinerja siswa. Kemudian pada faktor atribut data personal siswa, Gender atau jenis kelamin menjadi atribut yang paling sering digunakan pada

prediksi kinerja siswa. Dan pada faktor atribut kondisi sosial-ekonomi, parent occupation atau pekerjaan orang tua menjadi atribut yang paling sering digunakan pada prediksi kinerja siswa.

Selain metode dan atribut, dataset juga diperhatikan untuk diidentifikasi pada review ini. Dataset yang digunakan pada beberapa penelitian tentang prediksi kinerja siswa merupakan dataset privat yang berasal dari sekolah maupun universitas. Dan dataset universitas menjadi dataset yang paling sering digunakan dalam prediksi kinerja siswa karena disamping data yang banyak dan komplek juga terdapat banyak sekali celah dalam pengolahan data mahasiswa. Data mahasiswa yang terolah dengan baik juga dapat menentukan capaian mahasiswa di masa yang akan mendatang.

Pada review ini jauh dari kesempurnaan sehingga diharapkan review-review baru tentang prediksi kinerja siswa masih terus diteliti dan menjadi jauh lebih baik dari review-review yang telah ada. Untuk kedepannya masih ada celah dari review ini yang masih bisa menjadi objek penelitian diantara lain adalah memperbanyak studi literatur yang mana di review ini hanya berjumlah 35 studi. Kemudian menambah fokus pertanyaan penelitian dengan lebih rincinya secara detail atau tidak dari sisi umumnya saja.

Pada akhirnya, daftar studi utama disajikan pada Table 5. Daftar ini terdiri dari 5 atribut yaitu tahun, author, countries, methods, dan datasets dari 35 jurnal dan prosiding yang dirangkum berdasarkan tahun dari mulai 2010 hingga 2016.

Figure 16 menunjukkan peta pemikiran lengkap yang menyajikan hasil tinjauan literatur sistematis tentang prediksi kinerja siswa. Peta pemikiran telah digunakan untuk mengeksplorasi hubungan antara ide-ide dan unsur argumen dan untuk menghasilkan solusi untuk masalah. Hal ini membuat untuk lebih mudah secara logis mengatur informasi dan mengintegrasikan pengetahuan baru. Dalam review ini, peta pemikiran digunakan untuk menyajikan hasil tinjauan literatur sistematis tentang prediksi kinerja siswa.

## **DAFTAR PUSTAKA**

Adhatrao, K., Gaykar, A., Dhawan, A., Jha,



- R., Honrao, V. (2013) 'Predicting Students' Performance Using Id3 and C4.5 Classification Algorithms', *International Journal of Data Mining and Knowledge Management Process*, 3(5), pp. 39–52. doi: 10.5121/ijdkp.2013.3504.
- Angeline, D. M. D. (2013) 'Association Rule Generation for Student Performance Analysis using Apriori Algorithm', *The SIJ Transactions on Computer Science Engineering & its Applications (CSEA)*, 1(1), pp. p12--16.
- Arsad, P. M., Buniyamin, N. and Manan, J. A. (2013) 'A neural network students' performance prediction model (NNSPPM)', *Smart Instrumentation, Measurement and Applications (ICSIMA), 2013 IEEE International Conference on*, (November), pp. 1–5. doi: 10.1109/ICSIMA.2013.6717966.
- Badr El Din Ahmed, A. and Sayed Elaraby, I. (2014) 'Data Mining: A prediction for Student's Performance Using Classification Method', *World Journal of Computer Application and Technology*, 2(2), pp. 43–47. doi: 10.13189/wjcat.2014.020203.
- Bambrah, C. *et al.* (2014) 'Mining Association Rules in Student Assessment Data', *International Journal of Advanced Research in Computer and Communication Engineering*, 3(3), pp. 5340–5342. Available at: [http://www.ijarce.com/upload/2014/march/IJARCE6H s nirali Mining Association Rules.pdf](http://www.ijarce.com/upload/2014/march/IJARCE6H%20s%20nirali%20Mining%20Association%20Rules.pdf).
- Baradwaj, B. and Pal, S. (2012) 'Mining educational data to analyze student's performance', *International Journal of Advanced Computer Science and Applications*, 2(6), pp. 63–69. doi: vol.2,No.6.
- Bhardwaj, B. K. (2011) 'Data Mining: A prediction for performance improvement using classification', (*IJCSIS*) *International Journal of Computer Science and Information Security*, 9(4).
- Bisri, A. and Satria Wahono, R. (2015) 'Penerapan Adaboost untuk Penyelesaian Ketidakseimbangan Kelas pada Penentuan Kelulusan Mahasiswa dengan Metode Decision Tree', *Journal of Intelligent Systems*, 1(1), pp. 27–32.
- Budgen, D. and Brereton, P. (2006) 'Performing systematic literature reviews in software engineering', *Int. Conf. Soft. Engin.*, 45(4ve), p. 1051. doi: 10.1145/1134285.1134500.
- Bunkar, K. *et al.* (2012) 'Data mining: Prediction for performance improvement of graduate students using classification', *2012 Ninth International Conference on Wireless and Optical Communications Networks (WOCN)*, pp. 1–5. doi: 10.1109/WOCN.2012.6335530.
- Christian, T. M. and Ayub, M. (2014) 'Exploration of classification using NBTree for predicting students' performance', *Proceedings of 2014 International Conference on Data and Software Engineering, ICODSE 2014*, pp. 0–5. doi: 10.1109/ICODSE.2014.7062654.
- Echegaray-Calderon, O A. Barrios-Aranibar, D. (2015) 'Optimal selection of factors using Genetic Algorithms and Neural Networks for the prediction of students' academic performance', *2015 Latin America Congress on Computational Intelligence (LA-CCI)*, pp. 1–6. doi: 10.1109/LA-CCI.2015.7435976.
- Fu, J. H. *et al.* (2012) 'A support vector regression-based prediction of students' school performance', *Proceedings - 2012 International Symposium on Computer, Consumer and Control, IS3C 2012*, pp. 84–87. doi: 10.1109/IS3C.2012.31.
- Guleria, P. and Sood, M. (2015) 'Predicting Student Placements Using Bayesian Classification', pp. 109–112.
- Hamsa, Hashmia., Indiradevi, Simi., Kizhakkethottam, J. J. (2016) 'Student Academic Performance Prediction Model Using Decision Tree and Fuzzy Genetic Algorithm', *Procedia Technology*, 25, pp. 326–332. doi: 10.1016/j.protcy.2016.08.114.
- Jananto, A. (2013) 'Algoritma Naive Bayes untuk Mencari Perkiraan Waktu Studi Mahasiswa', *Teknologi Informasi DINAMIK*, 18(1), pp. 9–16.
- Jishan, S. T. *et al.* (2015) 'Improving accuracy of students' final grade

- prediction model using optimal equal width binning and synthetic minority over-sampling technique', *Decision Analytics*, 2(1), pp. 1–25. doi: 10.1186/s40165-014-0010-2.
- Kabakchieva, D. (2013) 'Predicting student performance by using data mining methods for classification', *Cybernetics and Information Technologies*, 13(1), pp. 61–72. doi: 10.2478/cait-2013-0006.
- Kaur, Kamaljit and Kaur, Kuljit (2015) 'Analyzing the effect of difficulty level of a course on students performance prediction using data mining', *2015 1st International Conference on Next Generation Computing Technologies (NGCT)*, (September), pp. 756–761. doi: 10.1109/NGCT.2015.7375222.
- Kitchenham, B. *et al.* (2009) 'Systematic literature reviews in software engineering - A systematic literature review', *Information and Software Technology*. Elsevier B.V., 51(1), pp. 7–15. doi: 10.1016/j.infsof.2008.09.009.
- Lopez Guarin, C. E., Guzman, E. L. and Gonzalez, F. A. (2015) 'A Model to Predict Low Academic Performance at a Specific Enrollment Using Data Mining', *Revista Iberoamericana de Tecnologias del Aprendizaje*, 10(3), pp. 119–125. doi: 10.1109/RITA.2015.2452632.
- Mehboob, B., Liaqat, R. M. and Saqib, N. A. (2016) 'Predicting Student Performance and Risk Analysis by Using Data Mining Approach', *International Journal of Computer Science and Information Security (IJCSIS)*, 14(7), pp. 69–76.
- Mishra, T., Kumar, D. and Gupta, S. (2014) 'Mining students' data for prediction performance', *International Conference on Advanced Computing and Communication Technologies, ACCT*, pp. 255–262. doi: 10.1109/ACCT.2014.105.
- Osmanbegović, E. and Suljić, M. (2012) 'Data mining approach for predicting student performance', *Journal of Economics and Business*, X(1), pp. 3–12.
- P. Pathak, N. Bansal, S. S. (2015) 'Mulyankan: A prediction for student's performance using Neural Network', *Computing for Sustainable Global Development (INDIACom)*, 2015 2nd International Conference on, pp. 46–49.
- Prachuabsupakij, W. and Soonthornphisaj, N. (2014) 'Hybrid sampling for multiclass imbalanced problem: Case study of students' performance prediction', *Advanced Computer Science and Information Systems (ICACSIS)*, 2014 International Conference on, pp. 321–326.
- Pratiwi, O. N. (2013) 'Predicting student placement class using data mining', *Proceedings of 2013 IEEE International Conference on Teaching, Assessment and Learning for Engineering, TALE 2013*, (August), pp. 618–621. doi: 10.1109/TALE.2013.6654511.
- Quadri, M. and Kalyankar, D. (2010) 'Drop out feature of student data for academic performance using decision tree techniques', *Global Journal of Computer*, 10(2), pp. 2–5. Available at: <http://computerresearch.org/stpr/index.php/gjcst/article/viewArticle/128>.
- Rahmani, B. and Aprilianto, H. (2014) 'Early Model of Student's Graduation Prediction Based on Neural Network', *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 12(2), p. 465. doi: 10.12928/TELKOMNIKA.v12i2.1603.
- Ramaswami, M. (2012) 'Student Performance Prediction', *International Journal of Computational Intelligence and Informatics*, 1(4), pp. 231–235.
- Ramesh, V., Parkavi, P. and Ramar, K. (2013) 'Predicting Student Performance : A Statistical and Data Mining Approach', *International Journal of Computer Applications*, 63(8), pp. 35–39. doi: 10.1504/IJTEL.2012.051816.
- Rashu, R. I., Haq, N. and Rahman, R. M. (2014) 'Data mining approaches to predict final grade by overcoming class imbalance problem', *2014 17th International Conference on Computer and Information Technology, ICCIT 2014*, pp. 14–19. doi: 10.1109/ICCITechn.2014.7073095.
- Ridwan, M., Suyono, H. and Sarosa, M. (2013) 'Penerapan Data Mining Untuk Evaluasi Kinerja Akademik Mahasiswa Menggunakan Algoritma Naive Bayes

- Classifier', *Eeccis*, 7(1), pp. 59–64.
- Shahiri, A. M., Husain, W. and Rashid, N. A. (2015) 'A Review on Predicting Student's Performance Using Data Mining Techniques', *Procedia Computer Science*. Elsevier Masson SAS, 72, pp. 414–422. doi: 10.1016/j.procs.2015.12.157.
- Sharabiani, A. *et al.* (2014) 'An Enhanced Bayesian Network Model for Prediction of Students' Academic Performance in Engineering Programs', *Proceedings of the IEEE Global Engineering Education Conference (EDUCON)*, (April), pp. 832–837. doi: 10.1109/EDUCON.2014.6826192.
- Tahyudin, I., Utami, E. and Amborowati, A. (2013) 'Comparing Clasification Algorithm Of Data Mining to Predict the Graduation Students on Time', *Information Systems International Conference (ISICO)*, (December), pp. 2–4.
- Tair, M. M. A. and El-Halees, A. M. (2012) 'Mining Educational Data to Improve Students' Performance: A Case Study', *International Journal of Information and Communication Technology Research*, 2(2), pp. 140–146.
- Yadav, S., Bharadwaj, B. and Pal, S. (2012) 'Data mining applications: A comparative study for predicting student's performance', *International Journal of Innovative Technology & Creative Engineering*, 1(12), pp. 13–19. Available at: <http://arxiv.org/abs/1202.4815>.
- Yadav, S. K. and Pal, S. (2012) 'Data Mining : A Prediction for Performance Improvement of Engineering Students using Classification', *World of Computer Science and Information Technology Journal WCSIT*, 2(2), pp. 51–56. doi: 10.1142/9789812771728\_0012.