

# Integrasi *Data Mining* dan *Online Analytical Processing* (OLAP) Pada Data Performa Siswa

Lydia Liliana<sup>[1]\*</sup>, Henny Hartono<sup>[2]</sup>, Devi Yurisca Bernanda<sup>[3]</sup>

Department of Information System <sup>[1], [2], [3]</sup>

Universitas Bunda Mulia

Jakarta, Indonesia

lydialiliana6@gmail.com<sup>[1]</sup>, hhartono@bundamulia.ac.id<sup>[2]</sup>, dbernanda@bundamulia.ac.id<sup>[3]</sup>

**Abstract**— Technological growth has an impact on increasing data. The accumulation of this data has created a pattern of more and more data, but the information obtained from this data is still minimal. Therefore, currently we need a data analysis technique in finding patterns from these data sets, namely data mining. The education sector produces various kinds of data, such as data on student performance in preparation for examinations. The data can be analyzed using the On-Line Analytical Processing (OLAP) method. This study focuses on the data mining integration process consisting of association, clustering, classification and forecasting with a combination of OLAP methods on student performance data using Power OLAP tools. The result of this research is the discovery of new patterns in the process of identifying groups of data, such as information about the average student exam results based on test preparation which is done in graphical form, so as to determine strategies in increasing student graduation. In further research, the author will collect student performance data derived from primary data by direct interviews with informants.

**Keywords**— *Data Mining, On-Line Analytical Processing, Power OLAP*

**Abstrak**—Pertumbuhan teknologi membawa dampak terhadap peningkatan data. Akumulasi data tersebut telah menciptakan pola data yang semakin banyak, namun perolehan informasi dari data tersebut masih minim. Oleh karena itu, saat ini diperlukan suatu teknik analisa data dalam mencari pola dari kumpulan data tersebut, yaitu data mining. Bidang pendidikan menghasilkan berbagai macam data, seperti data performa siswa dalam persiapan mengikuti ujian. Data tersebut dapat dianalisis dengan menggunakan metode *On-Line Analytical Processing* (OLAP). Penelitian ini berfokus pada proses integrasi data mining yang terdiri dari *association, clustering, classification* dan *forecasting* dengan kombinasi metode OLAP pada data performa siswa menggunakan *tools Power OLAP*. Hasil dari penelitian ini adalah penemuan pola baru dalam proses identifikasi kelompok data, seperti informasi mengenai rata-rata hasil ujian siswa berdasarkan persiapan ujian yang dilakukan dalam bentuk grafik, sehingga dapat menentukan strategi dalam meningkatkan kelulusan siswa. Pada penelitian selanjutnya, penulis akan melakukan pengumpulan data performa siswa yang berasal dari data primer dengan wawancara langsung ke narasumber.

**Kata Kunci**—*Data Mining, On-Line Analytical Processing, Power OLAP*

## I. PENDAHULUAN

Penggunaan teknologi informasi dan sistem informasi sudah menjangkau hampir segala bidang dalam aktivitas kehidupan manusia [1], [2]. Dukungan perkembangan teknologi yang semakin tinggi, maka akan membawa peningkatan terhadap kemampuan dalam melakukan pengumpulan serta mengolah data menjadi suatu informasi yang berguna [3], [4]. Akumulasi dari banyaknya data-data tersebut telah menciptakan pola dimana data yang semakin banyak namun perolehan informasi dari data tersebut masih minim [5], [6]. Oleh karena itu, saat ini diperlukan suatu teknik analisa data dalam mencari pola dari kumpulan data-data tersebut, seperti penggunaan *data mining* [7].

*Data mining* merupakan suatu teknik penemuan pola atau pengetahuan dari suatu kumpulan data yang banyak [8], [9]. Pola-pola tersebut dapat dilihat sebagai ringkasan dari *input* data yang dapat digunakan dalam bidang *On-Line Analytical Processing* (OLAP) untuk mendapatkan informasi yang berkualitas serta menemukan pola dari sejumlah besar data, sehingga didapatkan hasil dari kecenderungan data tersebut [10], [11]. Oleh karena itu, proses integrasi *data mining* yang dikombinasikan dengan teknik OLAP perlu dilakukan untuk membantu seseorang dalam melakukan pengambilan keputusan dan menetapkan berbagai kebijakan [12]. Sejauh ini penggunaan *data mining* dapat diterapkan pada beberapa bidang, salah satu contohnya dalam bidang pendidikan. Pada dunia pendidikan, data siswa dapat menghasilkan informasi yang berlimpah, seperti informasi performa siswa dalam persiapan mengikuti ujian.

Berdasarkan uraian tersebut, pada penelitian ini penulis melakukan pengolahan data performa siswa dalam persiapan mengikuti ujian dengan menggunakan 4 metode *data mining* yang dikombinasikan dengan metode *On-Line Analytical Processing* (OLAP). Data performa siswa yang digunakan merupakan data pada tahun 2018. Pengolahan data performa siswa ini perlu dilakukan untuk mengatasi permasalahan yang dialami berupa berlimpahnya data performa siswa namun makna atau informasi yang diperoleh dari data tersebut belum diketahui, sehingga institusi pendidikan merasa kebingungan dalam menentukan strategi yang tepat untuk meningkatkan performa siswa mereka. Metode *data mining* yang digunakan dalam penelitian ini adalah *association, clustering, classification* dan *forecasting*. Proses pengolahan data ini juga menggunakan bantuan *Power OLAP* yang merupakan salah

satu *tools business intelligence* dalam mengelola dan memberikan jawaban *realtime* sebagai dukungan dalam membantu pengambilan keputusan.

Penggunaan *tools Power OLAP* yang dikombinasikan dengan 4 metode *data mining* menghasilkan pola baru dalam proses identifikasi kelompok dalam data tersebut. Penelitian ini bertujuan untuk melakukan pengolahan terhadap data siswa agar didapat informasi penting berupa pengetahuan baru, seperti informasi mengenai rata-rata hasil ujian siswa berdasarkan proses persiapan ujian yang telah dilakukan dalam bentuk grafik sebagai alat pemodelan dari data, sehingga pengetahuan baru tersebut dapat membantu pihak universitas/sekolah untuk melakukan klasifikasi mengenai tingkat kelulusan dan dapat menentukan strategi dalam meningkatkan kelulusan siswa pada tahun-tahun berikutnya.

## II. KAJIAN LITERATUR

### A. Data Mining

*Data mining* merupakan proses pengumpulan informasi penting dari suatu kumpulan data besar [13]. Penelitian sebelumnya yang berjudul “*Model Business Intelligence System Design of Quality Products by Using Data Mining in R Bakery Company*” oleh R Fitriana, J. Saragih dan N. Luthfiana menjabarkan model *business intelligence* yang dibuat dengan proses *Extract Transform Load (ETL)* dan dikombinasikan dengan model *data mining* untuk meningkatkan kualitas produk di pabrik roti. Hasil yang diperoleh dari penelitian ini adalah jumlah roti yang tidak layak jual selama periode 7 bulan sebesar 96.744 buah berdasarkan hasil *OLAP Cube Method* [14]. Penggunaan *data mining* dapat menemukan pola tertentu yang tersembunyi, sehingga informasi yang dihasilkan dapat berbeda-beda dan unik. Metode *data mining* terdiri dari 4 hal, yaitu *association*, *clustering*, *classification* dan *forecasting* [15], [16].

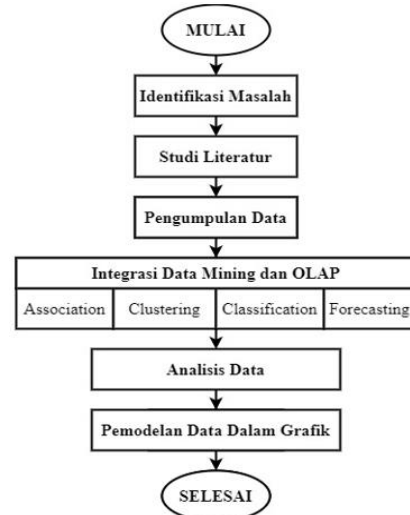
### B. On-Line Analytical Processing (OLAP)

*OLAP (On-Line Analytical Processing)* merupakan metode analisis data untuk menghasilkan laporan informasi yang dapat berguna bagi pengguna [17]. Penelitian sebelumnya yang berjudul “*On-Line Analytic Processing (OLAP) Modeling For Graduation Data Presentation*” oleh Arik Sofan Tohir, Kusri dan Sudarmawan menjabarkan pemodelan presentasi data wisuda dengan hasil data lulusan dua prodi dari *data warehouse* [18]. *OLAP* berfokus pada penyediaan jawaban terhadap *query* analisis multidimensi secara cepat [19], [20]. Struktur operasional utama pada *OLAP* didasarkan pada konsep “*Cube*”. *Cube* dalam *OLAP* merupakan struktur data multidimensial (aktual/virtual) yang memungkinkan analisis data secara cepat [21], [22]. Sistem *OLAP* mempunyai 2 kategori, yaitu *OLAP cube* dan *OLAP access tools* [23], [24].

## III. METODE PENELITIAN

Berdasarkan gambar 1, dapat diketahui bahwa penelitian ini dimulai dari [25], [26]: Identifikasi masalah, data performa siswa yang dihasilkan tiap institusi pendidikan sangat berlimpah yang mempunyai makna atau informasi tersembunyi dengan cara melakukan pengolahan data tersebut menggunakan teknik data mining. Studi literatur, melakukan pengumpulan referensi pustaka, seperti jurnal dan buku yang berhubungan

dengan *data mining* dan *OLAP*. Pengumpulan data, penelitian ini menggunakan data sekunder berupa data performa siswa dalam persiapan mengikuti ujian pada tahun 2018 yang diperoleh dari situs resmi *kaggle.com*. Integrasi *Data Mining* dan *OLAP* menggunakan bantuan *tools Power OLAP* dengan menganalisa data performa tersebut menggunakan 4 metode *data mining*, yaitu *association*, *clustering*, *classification* dan *forecasting*. Analisis data, menganalisa temuan pengetahuan baru yang didapatkan dari proses yang telah dilakukan sebelumnya. Pemodelan data dalam grafik dari metode *association*, *clustering*, *classification* dan *forecasting*.



Gambar 1. Metodologi Penelitian [27]

## IV. HASIL DAN PEMBAHASAN

Dari data performa siswa dalam persiapan mengikuti ujian, tersebut terdiri dari 9 atribut, yaitu *serial number*, *gender*, *race/ethnicity*, *parental level of education*, *lunch*, *Test preparation course*, *math score*, *writing score* dan *reading score*. Tabel 1 menunjukkan kolom *serial number* merupakan nomor unik yang menandakan urutan dari setiap data yang didapatkan. *Serial number* mempunyai rentang dari angka 1 sampai 1000. *Gender* yang digunakan adalah *female* (Perempuan) dan *male* (Laki-laki). Kolom *Race/Ethnicity* dikelompokkan berdasarkan suku-suku besar di Indonesia. *Group A* merupakan suku Chinese, *Group B* merupakan suku Jawa, *Group C* merupakan suku Sunda, *Group D* merupakan suku Betawi dan *Group E* merupakan suku Dayak. Kolom *parental level of education* memiliki arti tingkat pendidikan yang dicapai. *Associate's degree* setara dengan tingkat D3, *bachelor's degree* tingkat sarjana, *high school* tingkat SMA, *master's degree* tingkat S2, *some college* beberapa universitas, dan *some high school* beberapa sekolah SMA. Kolom *lunch* berarti harga makanan untuk setiap subyek, yaitu porsi *standard* dan *free/reduced* (Bisa mendapatkan harga makanan gratis). Kolom *test preparation course* menjelaskan mengenai latihan persiapan yang dilakukan subyek sebelum melaksanakan tes, yaitu persiapan telah selesai (*Completed*) dan belum ada persiapan (*None*). Kemudian pada data ini juga dilampirkan data *math score*, *reading score* dan *writing score*. Penelitian ini menggunakan 4 metode *data mining*, yaitu *association*, *clustering*, *classification* dan *forecasting*.

TABEL I. CONTOH DATA PERFORMA SISWA

Serial Number	Gender	Race/Ethnicity	Parental Level of Education	Lunch	Test Preparation Course	Math Score	Reading Score	Writing Score
1	Male	Group A	Associate's Degree	Free/Reduced	None	47	57	44
2	Female	Group B	Bachelor's Degree	Standard	None	72	72	74
3	Female	Group B	Master's Degree	Standard	None	90	95	93
4	Female	Group C	Some College	Standard	Completed	69	90	88
5	Male	Group D	High School	Free/Reduced	Completed	64	64	67
6	Male	Group E	Some High School	Standard	None	30	26	22

A. Association

Association merupakan proses untuk menemukan suatu kombinasi item dalam suatu waktu. Input yang digunakan berupa gender, test preparation course, serta math score, reading score dan writing score. Proses association ini membandingkan gender dengan test preparation course dan jumlah rata-rata nilai paling tinggi. Output yang dihasilkan adalah rata-rata nilai paling tinggi berdasarkan gender dan test preparation course. Proses pengolahan association ini menggunakan tools Power OLAP dengan langkah-langkah sebagai berikut:

- a) OLAP Basics. Membuat database baru sebagai tempat penyimpanan data performa siswa dan dimensi yang berisi atribut-atribut yang akan digunakan pada metode association, yaitu category, gender, serial number dan test course.
- b) OLAP Cube Modelling. Membuat cube baru bernama association. Cube ini akan diisi oleh atribut yang sebelumnya telah dipindahkan ke dimensi, yaitu category, gender, serial number dan test course. Penulis

tidak menggunakan rumus karena output yang dihasilkan adalah membandingkan dan mengetahui total score yang paling tinggi terdapat di gender apa dan sudah melakukan persiapan tes atau belum.

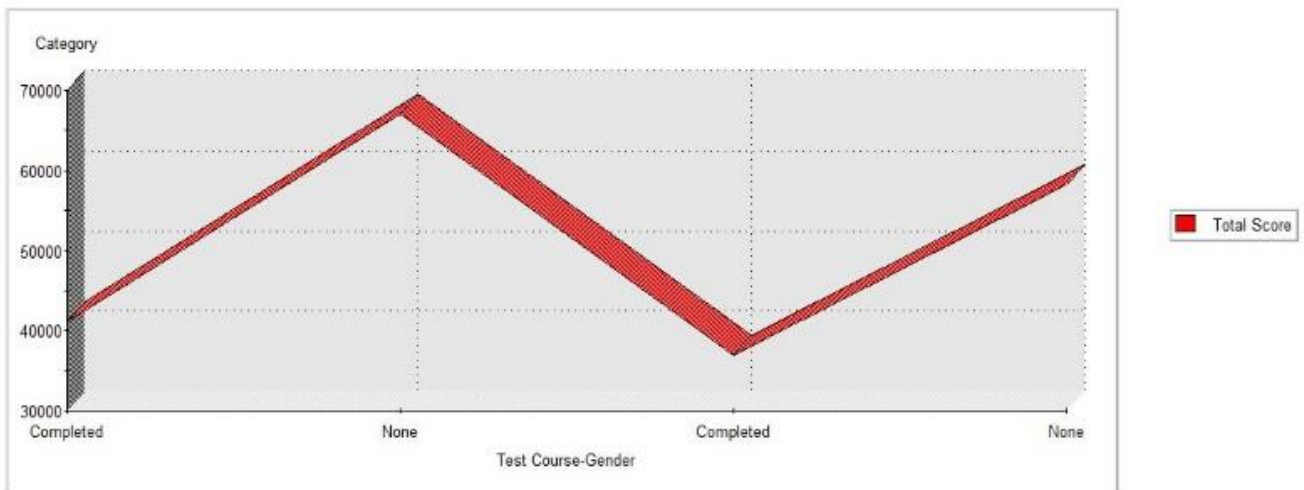
- c) OLAP Reporting. Membuat slice untuk mengatur data dalam cube. Setelah membuat slice baru, maka dapat membuka cube association yang telah dibuat, sehingga didapatkan hasil sebagai berikut:

TABEL II. HASIL SLICE ASSOCIATION

	Female		Male	
	Completed	None	Completed	None
Total Score	41099.00	67012.00	36948.00	58253.00

Berdasarkan tabel II, total score yang didapat dari masing-masing gender merupakan hasil dari penjumlahan keseluruhan math score, reading score dan writing score baik dari siswa perempuan maupun laki-laki.

No



Gambar 2. Grafik Hasil Association

Gambar 2 menunjukkan bahwa total score paling tinggi terdapat di jenis kelamin perempuan yang dimana belum mengikuti latihan persiapan untuk test. Sedangkan, total score paling rendah didapatkan dari jenis kelamin laki-laki yang

sudah mengikuti latihan persiapan untuk test. Maka dari itu, ditemukan suatu pola bahwa tidak selamanya seseorang yang sudah melakukan persiapan mendapatkan nilai terbaik. Hal ini terbukti pada hasil gender laki-laki yang sudah melakukan

persiapan justru mendapatkan *score* yang kurang baik. Oleh karena itu, pihak institusi perlu menganalisa faktor-faktor apa saja yang mempengaruhi siswa tersebut saat melakukan latihan persiapan, seperti apakah siswa tersebut kesusahan dalam menghadapi materi, sehingga perlu melakukan pengajaran secara *private*/pribadi. Temuan ini membuat institusi pendidikan harus menyiapkan strategi dalam mengatur cara persiapan latihan tes bagi siswa laki-laki.

**B. Clustering**

*Clustering* adalah proses pengelompokkan sejumlah data/obyek ke dalam kelompok data sehingga setiap kelompok mempunyai data yang mirip. *Input* yang diperlukan adalah *race/ethnicity*, *math score*, *reading score* dan *writing score*. Proses *clustering* ini membandingkan *total score* dari *race/ethnicity* yang diperoleh *math score*, *reading score* dan *writing score*. *Output* yang diinginkan adalah *total score* dari *writing score* masing-masing *group*. Proses pengolahan *clustering* ini dilakukan dengan menggunakan *tools Power OLAP* dengan langkah-langkah sebagai berikut:

- a) *OLAP Basics*. Membuka *database* yang telah dibuat sebelumnya dan membuat dimensi yang berisi atribut-atribut yang akan digunakan pada metode *clustering*, yaitu *category*, *serial number* dan *group*.
- b) *OLAP Cube Modelling*. Membuat *cube* baru bernama *clustering* yang diisi oleh atribut berupa *category*, *serial*

*number* dan *group*. Penulis menggunakan rumus sebagai berikut:

$$All\ and\ \{“Category.\ Average”\}\ and\ \{“Group.\ Group\ A”\} = “Clustering”.\{“Category.\ Total\ Score”,\ “Group.\ Group\ A”\}/3;$$

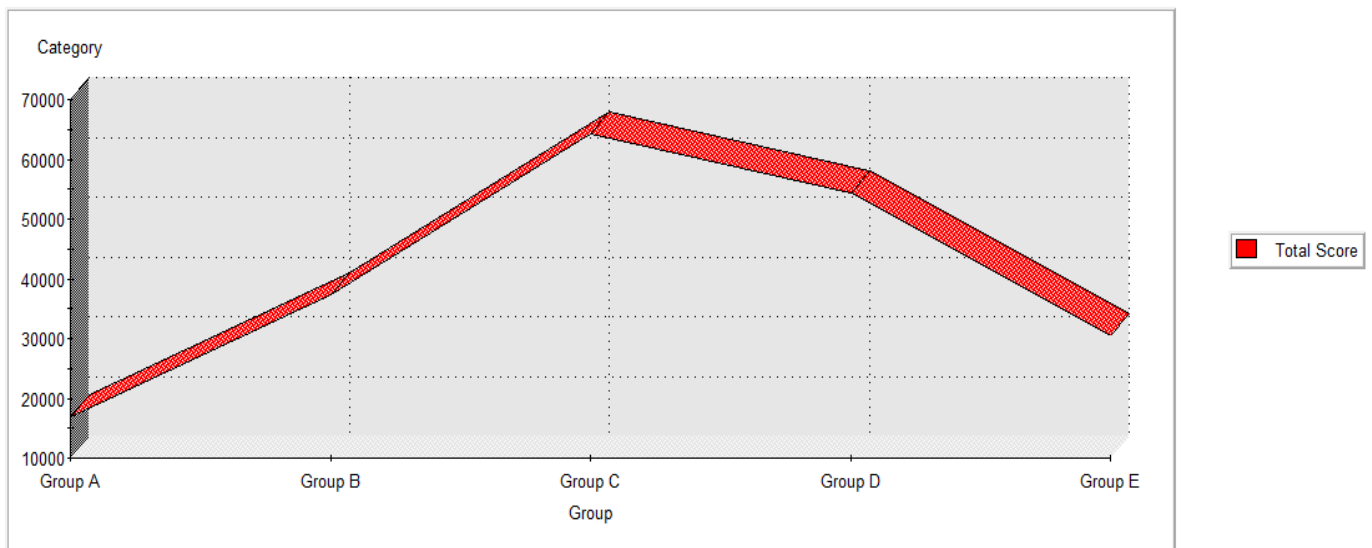
Rumus ini digunakan untuk menghitung *total score* dari *group A* yang berasal dari hasil penjumlahan *math score*, *reading score* dan *writing score* kemudian dibagi dengan 3. Angka 3 merupakan jumlah mata pelajaran yang ditempuh siswa tersebut. Rumus tersebut juga berlaku pada perhitungan *total score group B, C, D* dan *E*.

- c) *OLAP Reporting*. Membuat *slice* baru, maka dapat membuka *cube clustering* yang telah dibuat. Tabel 3 menunjukkan hasil dari *total score* masing-masing *group*.

TABEL III. HASIL SLICE CLUSTERING

	Total Score
Group A	16819.00
Group B	37317.00
Group C	64245.00
Group D	54375.00
Group E	30556.00

No



Gambar 3. Grafik Hasil Clustering

Gambar 3 menunjukkan bahwa *total score* yang paling tinggi terdapat di *group C* dan *total score* terendah terdapat di *group A*. Berdasarkan hasil tersebut, maka institusi pendidikan dapat menentukan strategi untuk meningkatkan pola pembelajaran siswa yang terdapat di dalam *group A*, seperti dengan melakukan penyesuaian lingkungan belajar siswa tersebut agar siswa terpacu untuk melakukan kegiatan tes persiapan sebelum

mengikuti ujian. Dengan menggunakan strategi tersebut, maka institusi pendidikan dapat mencapai target agar siswa di *group* tersebut dapat lulus pada tiap mata pelajaran.

**C. Classification**

*Classification* merupakan proses penemuan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas

data. *Input* yang digunakan adalah *parental level of education*. Proses *classification* ini membandingkan *parental level of education* dengan rata-rata penjumlahan *math score* dan *reading score* dari *parental level of education*. *Output* yang dihasilkan adalah rata-rata penjumlahan *math score* dan *reading score* dari masing-masing tingkatan kelulusan. Proses pengolahan *classification* ini dilakukan dengan menggunakan *tools Power OLAP* dengan langkah-langkah sebagai berikut:

- a) *OLAP Basics*. Membuka *database* yang telah dibuat dan membuat dimensi yang berisi atribut-atribut yang akan digunakan pada metode *classification*, yaitu *category*, *serial number* dan *parental level of education*.
- b) *OLAP Cube Modelling*. Membuat *cube* baru bernama *classification* yang diisi oleh atribut berupa *category*, *serial number* dan *parental level of education*. Penulis menggunakan rumus sebagai berikut:

$$\text{All and \{“Category. Average”\} = “Classification”.\{Category. Total Score\}};$$

Formula tersebut digunakan pada *Power OLAP* untuk membuat perhitungan rata-rata penjumlahan *math score* dan *reading score* dari 6 tingkatan yang ada.

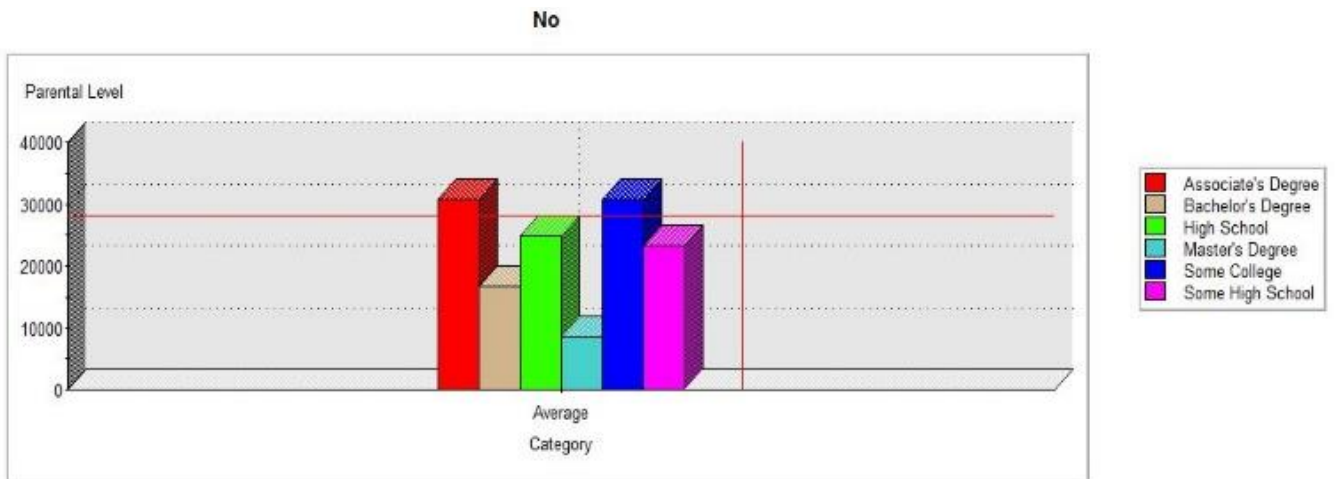
- c) *OLAP Reporting*. Tahapan selanjutnya adalah membuat *slice*. Setelah membuat *slice* baru, maka dapat membuka *cube classification* yang telah dibuat sebelumnya.

TABEL IV. HASIL SLICE CLASSIFICATION

	Average
Associate's Degree	30816.00
Bachelor's Degree	16802.00
High School	24861.00
Master's Degree	8562.00
Some College	30869.00
Some High School	23348.00

Tabel 4 menunjukkan hasil dari *slice classification*. *Associate's degree* rata-ratanya adalah 30816.00. *Bachelor's degree* rata-ratanya adalah 16802.00. *High school* rata-ratanya adalah 24861.00. *Master's degree* rata-ratanya adalah 8562.00. *Some college* rata-ratanya adalah 30869.00. *Some high school* rata-ratanya adalah 23348.00.

Gambar 4 menunjukkan bahwa *average* paling tinggi terdapat di lulusan D3 sedangkan *average* terendah terdapat di lulusan *master*. Dari data tersebut, membuktikan bahwa semakin tinggi tingkat lulusan seseorang, maka akan semakin sulit pula dalam melakukan proses pembelajaran untuk memperoleh nilai terbaik. Oleh karena itu, institusi pendidikan perlu menyusun strategi untuk menciptakan cara pembelajaran baru bagi para lulusan *master* supaya mudah dimengerti dengan mudah bagi siswa tersebut dan sebagai bahan persiapan untuk mengikuti ujian, sehingga hasil yang diperoleh dalam penilaian ujian akan lebih membaik.



Gambar 4. Grafik Hasil Classification

**D. Forecasting**

*Forecasting* adalah proses pengestimasi nilai prediksi berdasarkan pola-pola di dalam sekumpulan data. *Input* yang digunakan adalah *race/ethnicity* dan *average* untuk mencari rata-rata. Proses *forecasting* ini adalah data *race/ethnicity* dengan data jumlah dari rata-rata per orang dapat digunakan untuk memprediksi berapa kira-kira rata-rata tertinggi di tiap grup siswa di masa mendatang. *Output* yang dihasilkan adalah rata-rata dari masing-masing *group*. Proses pengolahan

*forecasting* ini dilakukan dengan menggunakan *tools Power OLAP* dengan langkah-langkah sebagai berikut:

- a) *OLAP Basics*. Membuka *database* yang dibuat sebelumnya dan membuat dimensi yang berisi atribut-atribut pada metode *forecasting*, yaitu *category*, *serial number* dan *group*.
- b) *OLAP Cube Modelling*. Membuat *cube* baru bernama *forecasting* yang diisi oleh atribut *category*, *serial*

number dan group. Penulis menggunakan rumus sebagai berikut:

$$All\ and\ \{“Category.\ Average”\}\ and\ \{“Group.\ Group\ A”\} = “Clustering”.\{“Category.\ Total\ Score”, “Group.\ Group\ A”\}/3;$$

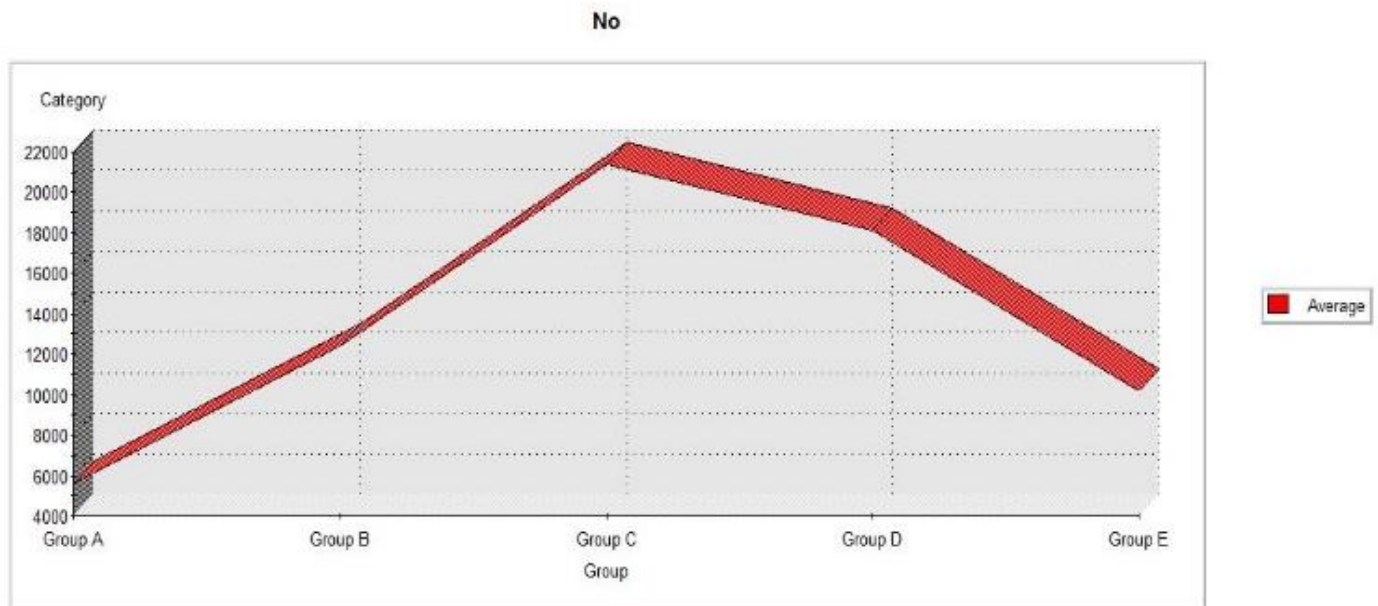
Rumus ini digunakan untuk menghitung *average* dari *group A* yang berasal dari hasil penjumlahan *math score*, *reading score* dan *writing score* kemudian dibagi dengan 3. Angka 3 merupakan jumlah mata pelajaran yang ditempuh siswa tersebut. Rumus tersebut juga berlaku pada perhitungan *average group B, C, D* dan *E*.

- c) *OLAP Reporting*. Membuat *slice* dan membuka *cube forecasting* yang telah dibuat sebelumnya. Tabel 5 menunjukkan hasil dari *slice forecasting* berupa rata-rata nilai *race/ethnicity*.

TABEL V. HASIL SLICE FORECASTING

	Average
Group A	5606.00
Group B	12439.00
Group C	21415.00
Group D	18125.00
Group E	10185.00

Gambar 5 menunjukkan bahwa *average* yang paling tinggi terdapat di *group C* dan *average* terendah terdapat di *group A*, sehingga dapat diprediksi bahwa *group C* di masa mendatang akan memiliki rata-rata yang semakin meningkat setiap periodenya. Oleh karena itu, institusi pendidikan dapat mengatur strategi untuk meningkatkan rata-rata perolehan nilai agar setiap *group* mendapatkan hasil yang terbaik. Selain itu, ada baiknya institusi dapat meramalkan atau memprediksi cara terbaik untuk meningkatkan nilai dari tiap siswa, seperti menambah ekstrakurikuler yang disesuaikan dengan bidang yang diminati.



Gambar 5. Grafik Hasil Forecasting

## V. KESIMPULAN

Kesimpulan yang diperoleh dari penelitian ini adalah Peningkatan penggunaan data telah menciptakan pola, dimana data yang semakin banyak namun perolehan informasi masih minim. Oleh karena itu, saat ini diperlukan suatu teknik analisa data, seperti *data mining* yang diintegrasikan dengan *On-Line Analytical Processing (OLAP)* untuk mendapatkan informasi yang berkualitas. Proses integrasi *data mining* dengan *OLAP* berfokus pada 4 metode *data mining*, berupa *association*, *clustering*, *classification* dan *forecasting*.

Metode *association* merujuk bahwa pihak institusi perlu menganalisa faktor-faktor dan strategi yang mempengaruhi siswa saat melakukan latihan persiapan. Metode *clustering* merujuk pada institusi pendidikan harus dapat menentukan

strategi untuk meningkatkan pola pembelajaran siswa agar mencapai target lulus.

Metode *classification* membuktikan bahwa semakin tinggi tingkat lulusan seseorang, maka akan semakin sulit dalam melakukan proses pembelajaran untuk memperoleh nilai terbaik. Metode *forecasting* merujuk ada baiknya institusi dapat memprediksi cara terbaik untuk meningkatkan nilai dari tiap siswa.

Secara keseluruhan, didapatkan pengetahuan baru untuk mengukur tingkat keberhasilan siswa dalam melakukan persiapan ujian dan sejauh mana institusi pendidikan sebagai lembaga penyelenggara dapat memfasilitasi siswa mereka dengan baik.

Pada penelitian selanjutnya, penulis akan melakukan pengumpulan data yang berasal dari data primer dengan wawancara langsung ke narasumber.

#### REFERENCES

- [1] M. Liu, M. Qu, and B. Zhao, "Research and Citation Analysis of Data Mining Technology Based on Bayes Algorithm," *Mob. Networks Appl.*, vol. 22, no. 3, pp. 418–426, 2017, doi: 10.1007/s11036-016-0797-2.
- [2] E. B. Lee, J. Kim, and S. G. Lee, "Predicting Customer Churn in Mobile Industry Using Data Mining Technology," *Ind. Manag. Data Syst.*, vol. 117, no. 1, pp. 90–109, 2017, doi: 10.1108/IMDS-12-2015-0509.
- [3] D. S. Susilawati and D. Riana, "Optimization the Naive Bayes Classifier Method to Diagnose Diabetes Mellitus," *IAIC Trans. Sustain. Digit. Innov.*, vol. 1, no. 1, pp. 78–86, 2019, doi: 10.34306/itsdi.v1i1.21.
- [4] B. G. Sudarsono, J. F. Andry, and Nirwan, "Design Information System Order Fulfillment Using Archimate Modelling," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 2, pp. 1303–1310, 2020, doi: 10.30534/ijatcse/2020/61922020.
- [5] H. Tannady, J. F. Andry, and F. Nurprihatin, "Determinants Factors Toward the Performance of the Employee in the Crude Palm Oil Industry in West Sumatera, Indonesia," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 771, pp. 1–6, 2020, doi: 10.1088/1757-899X/771/1/012066.
- [6] J. F. Andry, H. Tannady, and F. Nurprihatin, "Eliciting Requirements of Order Fulfillment in A Company," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 771, pp. 1–6, 2020, doi: 10.1088/1757-899X/771/1/012023.
- [7] W. Sun, Z. Cai, F. Liu, S. Fang, and G. Wang, "A Survey of Data Mining Technology on Electronic Medical Records," *2017 IEEE 19th Int. Conf. e-Health Networking, Appl. Serv. Heal. 2017*, pp. 1–6, 2017, doi: 10.1109/HealthCom.2017.8210774.
- [8] G. Singh and B. Kishore, "Comparative Analysis of Opinion Mining KDD for Customer Reviews," *Int. J. Res. Eng. Sci. Manag.*, vol. 2, no. 8, pp. 15–17, 2019.
- [9] F. A. Rahman, S. M. Shamsuddin, S. Hasan, and N. A. Haris, "A Review of KDD Data Mining Framework and Its Application in Logistics and Transportation," *Int. J. Supply Chain Manag.*, vol. 5, no. 2, pp. 77–84, 2016.
- [10] B. Salimi, C. Cole, P. Li, J. Gehrke, and D. Suciu, "HypDB: A Demonstration of Detecting, Explaining and Resolving Bias in OLAP Queries," *Proc. VLDB Endow.*, vol. 11, no. 12, pp. 2062–2065, 2018, doi: 10.14778/3229863.3236260.
- [11] C. Zhan *et al.*, "AnalyticDB: Real-Time OLAP Database System at Alibaba Cloud," *Proc. VLDB Endow.*, vol. 12, no. 12, pp. 2059–2070, 2018, doi: 10.14778/3352063.3352124.
- [12] P. Nahar and S. Kumar, "Developing a Decision Support System (DSS) for Utility of Power Supply Organization," *Int. J. Appl. Eng. Res.*, vol. 12, no. 9, pp. 1962–1968, 2017.
- [13] Z. Chen, "An Approach for Developing Platform of OLAP," *Int. Conf. Artif. Intell. Control Autom. Eng.*, pp. 117–121, 2019.
- [14] R. Fitriana, J. Saragih, and N. Luthfiana, "Model Business Intelligence System Design of Quality Products by Using Data Mining in R Bakery Company," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 277, no. 1, pp. 1–9, 2017, doi: 10.1088/1757-899X/277/1/012005.
- [15] F. Firat, A. K. Arslan, C. Colak, and H. Harputluoglu, "Estimation of Risk Factors Associated With Colorectal Cancer: An Application of Knowledge Discovery in Databases," *Kuwait J. Sci.*, vol. 43, no. 2, pp. 151–161, 2016.
- [16] A. Dutt, M. A. Ismail, and T. Herawan, "A Systematic Review on Educational Data Mining," *IEEE*, vol. 5, pp. 1–15, 2017.
- [17] M. Djedaini, K. Drushku, N. Labroche, P. Marcel, V. Peralta, and W. Verdeaux, "Automatic Assessment of Interactive OLAP Explorations," *Inf. Syst.*, vol. 82, pp. 148–163, 2019, doi: 10.1016/j.is.2018.06.008.
- [18] A. S. Tohir, K. Kusri, and S. Sudarmawan, "On-Line Analytic Processing (OLAP) Modeling for Graduation Data Presentation," *Proc. - 2017 2nd Int. Conf. Inf. Technol. Inf. Syst. Electr. Eng.*, pp. 132–135, 2018, doi: 10.1109/ICITISEE.2017.8285481.
- [19] A. Lamani, B. Erraha, M. Elkyal, and A. Sair, "Data Mining Techniques Application for Prediction in OLAP Cube," *Int. J. Electr. Comput. Eng.*, vol. 9, no. 3, pp. 2094–2102, 2019, doi: 10.11591/ijece.v9i3.pp2094-2102.
- [20] M. P. Cristescu, "Using OLAP Data Cubes in Business Intelligence," *Sci. Bull.*, vol. 21, no. 2, pp. 80–86, 2017, doi: 10.1515/bsaft-2016-0039.
- [21] R. Andreswari and M. A. Rasyid, "OLAP Cube Processing of Production Planning Real-life Event Log: A Case Study," *Atl. Highlights Eng.*, vol. 2, pp. 148–153, 2019.
- [22] N. Shaik, W. Ullah, and G. Pradeepni, "OLAP Cube-Based Graph Approach for Bibliographic Data," *Am. J. Eng. Res.*, vol. 5, no. 2, pp. 237–240, 2016.
- [23] Neha and K. Garg, "Efficient Information Retrieval Using Multidimensional OLAP Cube," *Int. Res. J. Eng. Technol.*, vol. 4, no. 7, pp. 2885–2889, 2017.
- [24] L. Gomez, B. Kuijpers, and A. Vaisman, "Performing OLAP Over Graph Data: Query Language, Implementation, and a Case Study," *ACM Int. Conf. Proceeding Ser.*, pp. 1–8, 2017, doi: 10.1145/3129292.3129293.
- [25] N. Shaik, W. Ullah, and G. Pradeepni, "OLAP Mining Rules: Association of OLAP with Data Mining," *Am. J. Eng. Res.*, vol. 5, no. 2, pp. 2320–847, 2016. [Online]. Available: www.ajer.org.
- [26] S. Hendrian, "Algoritma Klasifikasi Data Mining Untuk Memprediksi Siswa Dalam Memperoleh Bantuan Dana Pendidikan," *Fakt. Exacta*, vol. 11, no. 3, pp. 266–274, 2018, doi: 10.30998/faktorexacta.v11i3.2777.
- [27] H. Sulistiani and D. Darwis, "Penerapan Metode Agile untuk Pengembangan Online Analytical Processing (OLAP) pada Data Penjualan (Studi Kasus : CV Adilia Lestari)," *J. CoreIT*, vol. 6, no. 1, pp. 50–56, 2020.