Recommendation of the Best Majors in Higher Education for Students Career with Big Data

Inayatulloh
Information Systems Department
School of Information system
Bina Nusantara University
Jakarta, Indonesia 11480
Inay@ binus.ac.id

Rajiv Dharma Mangruwa
Communication and Business-Telkom University, Bandung, Indonesia
Azman Hashim International Business School –
University of Technology Malaysia , Malaysia
rmangruwa@telkomuniversity.ac.id

Rivaldhy N. Muhammad
Study Program of Law, Faculty of Law, Social and Political Sciences
Open University,Indonesia
valdhymuhammad1@gmail.com

Destiana Kumala
Sekolah Tinggi Ekonomi Bisnis Islam Syariah Bina Mandiri
Bogor, Indonesia 11480
destiana.kumala@binamandiri.ac.id

*Correspondent Author

Abstract

Higher education is an educational institution that provides educational services and offers many courses tailored to the needs of students and industry. The large number of skilled workforces needs in the industry requires higher education to prepare these skilled workers by building and offering various majors at universities. However, many students do not know the right major that suits their needs and desires that are aligned with industry needs. On the other hand, Big Data is a technology that collects very large amounts of data in real time which is then processed to produce important information that is needed by the user. exist in universities by using Big Data. The research method uses a qualitative approach through observation related to the difficulty in determining the choice of majors at the university. Literature studies are used to find alternative solutions with an information technology approach.

Keywords
Big Data, recommendation, higher education, career

1. Introduction
Higher education as an educational institution is a liaison between the workforce and industry. The high complexity of the company's business processes requires workers with several different areas of expertise, so workers with various

levels of education and areas of expertise are needed by the industry to occupy positions that will support the company's business processes (Mardis et al. 2018) (Borner et al. 2018) (Penprase 2018) .

Skilled workforce comes from students from schools with several levels of education where many students do not know the skills needed by companies that are in accordance with the interests, talents and abilities of students so that many students take the wrong major when studying in higher education (Syamsiah 2018) (Indira and Martha 2021) (Sohar 2018).

Meanwhile, Big data as a cutting-edge technology has the ability to process data and produce important information needed by users to provide solutions to student problems to determine the skills needed by the company/industry provided by higher education and tailored to the interests and talents of students. Thus, the purpose of this research is to help students or prospective workers choose the job and skills needed that are tailored to their interests and talents. The research method uses a qualitative approach with observations and interviews with several students related to the problem of determining careers in companies and interviews with academics in higher education to obtain information on majors offered by higher education. Literature studies are used to find alternative solutions with information technology.

2. Literature Review

Big Data

Big Data is a term used to define a very large number of structured and unstructured data sets. The data can later be processed for various purposes, especially technology and business needs (Dash et al. 2019) (Anshari et al 2019) (Mariani et al. 2018). The concept of Big Data has three important criteria that must be met by a data set to be included in the Big Data category, namely: Volume, where To meet the criteria as Big Data, the volume or size of data collected from various sources must be large. At least, the data collected is in units of terabytes (one trillion bytes) or exabytes (10^{18} bytes) to be categorized as Big Data . The second is the speed where in addition to the large volume, the speed of data collection in a matter of terabytes or exabytes must be able to be done in a short period of time. Even if the data volume criteria are met, if the data collection period is slow, the data set cannot be categorized as Big Data. The third is Variety. In addition to the criteria for size and speed, another criterion that must be met in order for a data set to be categorized as Big Data is the variety of data. The data collected must consist of various formats, such as structured data, unstructured data, text, video, audio, transaction lists, and various other formats. If the data consists of only a few formats, the data set cannot be categorized as Big Data. These three criteria must be met in order for a data set to be categorized as Big Data. If one of the criteria is not met then the data set cannot be categorized as Big Data (Elhoseny et al. 2018) (Mikalef et al. 2018) (Rabah 2018).

Big Data as a collection of data with large volumes, short collection times, and various data formats provide great benefits and potential for use. In the business world, processing Big Data properly can provide several advantages. For example, the broad scope of Big Data can provide in-depth analytical results that are difficult to obtain through conventional analytics in a short period of time. Later, the analysis results can be used to find new target markets, or determine appropriate promotional steps (Yang et al. 2020) (Tamimina et al. 2020) (Hajli et al. 2020).

The use of Big Data in the field of e-commerce is also not new. Online shopping sites provide recommendations for items that match your interests exactly, then you have found the application of Big Data in the e-commerce field. Not only in the business world, the technology world has also applied the use of Big Data to the development of artificial intelligence. Big Data is used by developers to determine various possible patterns of behavior under certain conditions (Husssein et al. 2021) (Atal 2020) (Shen et al. 2020).

3. Methods

Figure 1 explains the research method. This research is motivated by the confusion of school students to choose majors in higher education and careers in companies. The next stage after finding the problem, the research is continued with the root cause of the confusion that causes students to choose majors in higher education and careers in companies. After determining the root of the problem, the next step is to look for alternative solutions that can provide accurate information to students or prospective workers about talents, abilities of students and majors in higher education

related to careers that are still ahead. The final stage is to create a Big Data model to help students determine their majors and future careers.

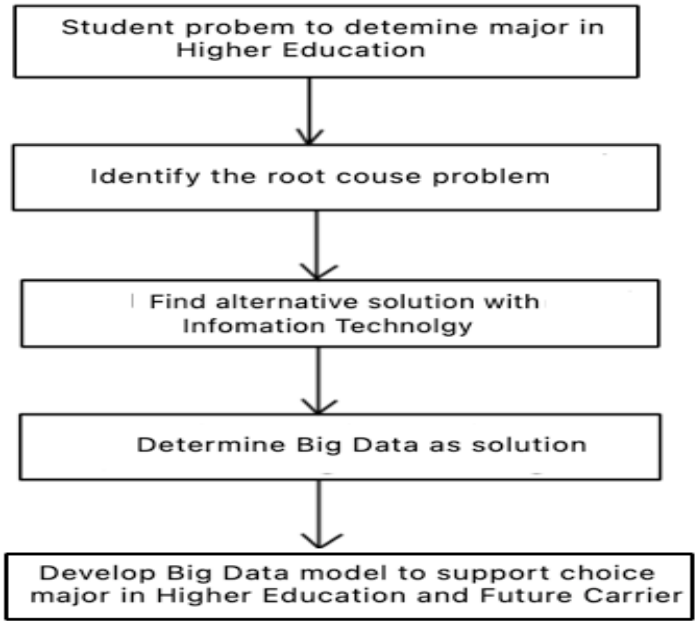

Figure 1. Research method.

4. Results and Discussion

Figure 2 show the proposes model. The proposed model consists of 3 parts of big data, namely:

a. Data Ingestion is the first layer in the Big Data Architecture, this is the layer that is responsible for collecting data from various data sources: IoT devices, data lakes, databases, and SaaS applications into a target data warehouse (Semlali et al. 2020) (Sharma et al. 2021) (Erraissi and Abdessamed 2018).

This is a critical point in the process because at this stage the size and complexity of the data can be understood, which will affect the architecture or every decision we make down the road. The function of Data ingestion layer is Availability. In this model the data is available to all users in universities, student, physician, lecturer, company, employee etc. a quality data ingestion process can turn different types of data into a unified data that is easy to read and perform statistics and manipulations on and data ingestion process saves engineers time in trying to collect the data they need and develop efficiently instead. Data ingestion in this model has several functions, namely:

i. Dynamic Data Exchange (DDE) is an inter proses communication system that allows data to be communicated or shared between applications within the operating system. Dynamic Data Exchange leverages shared memory and a set of commands, message formats, and protocols for communication and sharing. Object linking and embedding are successful in Dynamic Data Exchange, but the latter is still used for simple and easy inter process communication tasks. In this model, data exchange is used to communicate all processes from data in higher education, psychological institutions, company employee data and streaming data from the internet related to the required information (Schitt et al. 2018) (Zhang 2020) (Petter et al. 2018).

ii. Data synchronization is the process of establishing consistency between data from source to target data storage and vice versa and continuous harmonization of data over time. This is important for a wide variety of applications, including file synchronization and hardware synchronization. In the proposed model, synchronization of all data from higher education, psychological institutions,

company employee data and streaming data from the internet related to the required information is carried out in real time (Di et al. 2020) (Wang et al. 2018) (Wang et al. 2018).

iii. Data migration is the process of moving or transforming data from one context to another. The context can be in the form of data structures, data formats, technology platforms, or locations. In the proposed model, data migration from higher education, psychological institutions, company employee data and streaming data from the internet is converted according to the data structure, data format, technology platform, or location that has been determined (Mazumdar et al. 2019) (Hababeh et al. 2018) (Manogaran et al. 2020).

b. The second process is Storage and processing. This process includes data validation from higher education, psychology institutions, company employee data and streaming data from the internet. Then the data also goes through a clean process, normalization. The last process is curation. Data curation is the process of creating, organizing, and maintaining data sets so they can be accessed and used by people looking for information.

c. The result of the Big Data process is in the form of visualization of information that can be tailored to the needs of the user. Big data can predict and recommend students about careers, skills that are tailored to students' talents and abilities. Big data can also generate major recommendations that are in line with students' future career plans.

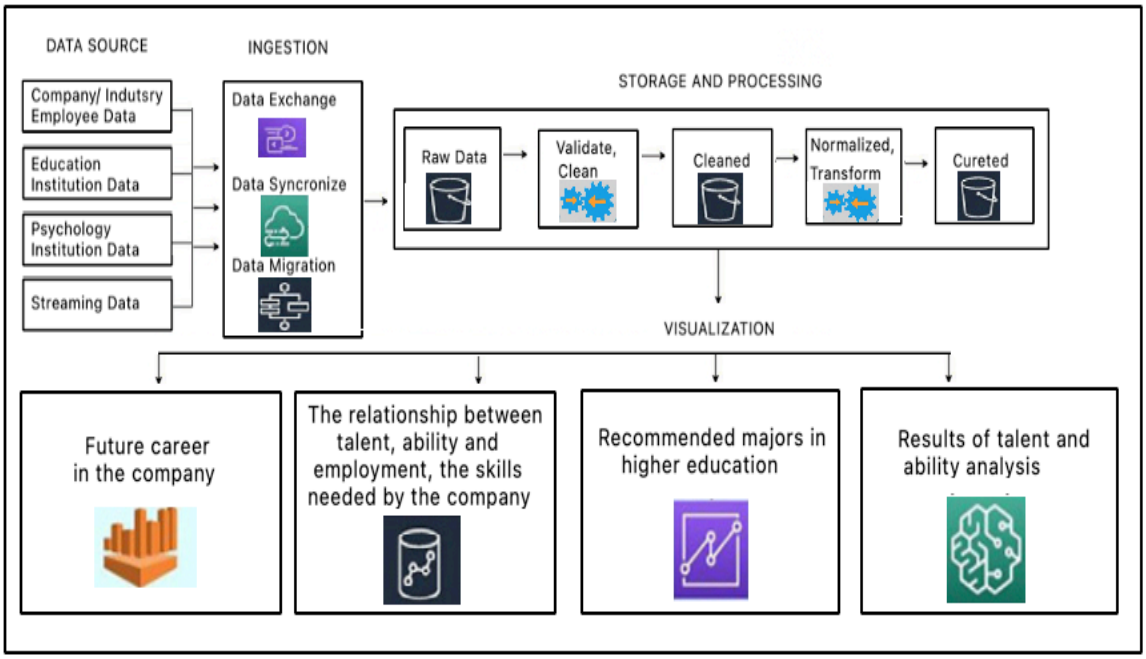

Figure 2. Big Data Model for Student Career Recommendation

5. Conclusion

The adoption of big data to provide recommendations for majors in higher education for students will solve the problem of students' confusion to plan their future. Combining data and information based on student psychology and talent, students' abilities with career information in the company and streaming data based on existing employee profiles will produce comprehensive information for students or prospective workers so that they can plan their careers in the future. Future career planning starts from determining the major in higher education that is in accordance with the talents and abilities of students.

References

Anshari, Muhammad, et al. "Customer relationship management and big data enabled: Personalization & customization of services." *Applied Computing and Informatics* 15.2 (2019): 94-101.

Atal, Nimita Shyamsunder. *Application Of Big Data Analytics Framework For Enhancing Customer Experience On E-Commerce Shopping Portals*. Diss. Purdue University Graduate School, 2020.

Börner, Katy, et al. "Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy." *Proceedings of the National Academy of Sciences* 115.50 (2018): 12630-12637.

Dash, Sabyasachi, et al. "Big data in healthcare: management, analysis and future prospects." *Journal of Big Data* 6.1 (2019): 1-25.

Di Leoni, Patricio Clark, Andrea Mazzino, and Luca Biferale. "Synchronization to big data: Nudging the Navier-Stokes equations for data assimilation of turbulent flows." *Physical Review X* 10.1 (2020): 011023.

Elhoseny, Mohamed, et al. "A hybrid model of internet of things and cloud computing to manage big data in health services applications." *Future generation computer systems* 86 (2018): 1383-1394.

Erraissi, Allae, and Abdessamad Belangour. "Data sources and ingestion big data layers: meta-modeling of key concepts and features." *International Journal of Engineering & Technology* 7.4 (2018): 3607-3612.

Hababeh, Ismail, et al. "An integrated methodology for big data classification and security for improving cloud systems data mobility." *IEEE Access* 7 (2018): 9153-9163.

Hajli, Nick, et al. "Understanding market agility for new product success with big data analytics." *Industrial Marketing Management* 86 (2020): 135-143.

Hussien, Farah Tawfiq Abdul, Abdul Monem S. Rahma, and Hala Bahjat Abdul Wahab. "Recommendation Systems For E-commerce Systems An Overview." *Journal of Physics: Conference Series*. Vol. 1897. No. 1. IOP Publishing, 2021.

Indira, Lydia, and Martha Ayu. " The Relationship of Authoritarian Parenting with Impostor Syndrome in Students in the Wrong Major." *Intensi: Jurnal Psikologi* 1.1 (2021).

Mariani, Marcello, et al. "Business intelligence and big data in hospitality and tourism: a systematic literature review." *International Journal of Contemporary Hospitality Management*(2018).

Manogaran, Gunasekaran, et al. "FDM: Fuzzy-optimized data management technique for improving big data analytics." *IEEE Transactions on Fuzzy Systems* 29.1 (2020): 177-185.

Mardis, Marcia A., et al. "Assessing alignment between information technology educational opportunities, professional requirements, and industry demands." *Education and Information Technologies* 23.4 (2018): 1547-1584.

Mikalef, Patrick, et al. "Big data analytics and firm performance: Findings from a mixed-method approach." *Journal of Business Research* 98 (2019): 261-276.

Mazumdar, Somnath, et al. "A survey on data storage and placement methodologies for cloud-big data ecosystem." *Journal of Big Data* 6.1 (2019): 1-37.

Potter, Anthony, et al. "Dynamic data exchange in distributed RDF stores." *IEEE Transactions on Knowledge and Data Engineering* 30.12 (2018): 2312-2325.

Penprase, Bryan Edward. "The fourth industrial revolution and higher education." *Higher education in the era of the fourth industrial revolution* 10 (2018): 978-981.

Rabah, Kefa. "Convergence of AI, IoT, big data and blockchain: a review." *The lake institute Journal* 1.1 (2018): 1-18.

Shen, Jian, Tianqi Zhou, and Lina Chen. "Collaborative filtering-based recommendation system for big data." *International Journal of Computational Science and Engineering* 21.2 (2020): 219-225.

Semlali, Badr-Eddine Boudriki, Chaker El Amrani, and Guadalupe Ortiz. "SAT-ETL-Integrator: an extract-transform-load software for satellite big data ingestion." *Journal of Applied Remote Sensing* 14.1 (2020): 018501.

Sharma, Garima, Vikas Tripathi, and Awadhesh Srivastava. "Recent trends in big data ingestion tools: a study." *Research in Intelligent and Computing in Engineering*. Springer, Singapore, 2021. 873-881.

Sohar, Aminullah Cik. " Reality Counseling Approach in Overcoming Intrapersonal Conflict in "Wrong Major" Students" . *Jurnal Komunikasi Islam dan Kehumasan (JKPI)* 2.1 (2018): 1-15.

Syamsiah, Ary Nur. *Stress coping strategies for students in the wrong major at three state universities in Malang City*. Diss. Universitas Negeri Malang, 2018.

Schmitt, Alexandre, Florent Carlier, and Valérie Renault. "Dynamic bridge generation for IoT data exchange via the MQTT protocol." *Procedia computer science* 130 (2018): 90-97.

Tamiminia, Haifa, et al. "Google Earth Engine for geo-big data applications: A meta-analysis and systematic review." *ISPRS Journal of Photogrammetry and Remote Sensing* 164 (2020): 152-170.

Wang, Yuan, et al. "Position synchronization for track geometry inspection data via big-data fusion and incremental learning." *Transportation Research Part C: Emerging Technologies* 93 (2018): 544-565.

Wang, Tian, et al. "Fog-based computing and storage offloading for data synchronization in IoT." *IEEE Internet of Things Journal*6.3 (2018): 4272-4282.

Yang, Jin, et al. "Brief introduction of medical database and data mining technology in big data era." *Journal of Evidence-Based Medicine* 13.1 (2020): 57-69.

Zhang, YuHong. "Rely on Negotiable Hash Algorithm of Data Exchange in the Relational Data Privacy Protection." *Journal of Physics: Conference Series*. Vol. 1533. No. 4. IOP Publishing, 2020.

Biographies

Inayatulloh SE.MMSI,CDMS.CSCA is a lecturer at Bina Nusantara University, School of Information System Jakarta Indonesia and also a doctoral candidate of computer science, experienced in managing systems in the retail, automotive, convection and education industries. research domain in e-learning, e-business, e-commerce, cloud computing, IoT and block chain technology.

Dr. Ir. Nico Djundharto Djajasinga, MSc. CPFF. IPM is a lecturer and head of the Applied Masters Study Program at the Indonesian Land Transportation Polytechnic-STTD and has competence as an examiner of railroad buildings, inspector of railway infrastructure, railway auditor and assessor of railway HR and also has expertise in the field of land transportation engineering, especially railways, service management and logistics, as well as safety and risk

Destiana Kumala, S.kom, M.M., C. HR. Adm. Completed Bachelor of Information Systems Education (2009) and Master of Management with a concentration on Human Resources (2012), besides that the author has also completed training in Human Resources Administration (2021). Currently active as a permanent lecturer at STEBIS Bina Mandiri Cileungsi in the sharia business management study program and focuses on the field of management science. Apart from being the vice chairman of the academic field, the author is also a researcher in the fields of economics, management and computers

Rivaldhy N. Muhammad is a Lecturer at the Open University which is one of the State Universities in Indonesia. holds a Bachelor of Laws degree at Khairun University Ternate and a Master's degree at Janabadra University Yogyakarta. Since becoming a Lecturer in 2018, he has been assigned to Jayapura Papua Open University as a Lecturer in the field of Legal Studies who always provides teaching and guidance to students who need good knowledge. Before becoming a lecturer, Rivaldhy N. Muhammad was a lawyer who always helped people who needed legal aid services.