ABSTRACT

Data integration is the process of combining data from various different data sources into one data source that can be used by a company. This is done so that all worker positions or divisions within a company can view the data as a whole without having to make requests to the relevant divisions to view the data. However, this process can cause problems with data confidentiality if the data cleaning and filtering process is not carried out. Therefore, the Capstone Design Project designed a data fabric and machine learning system that was implemented on a website to increase revenue and scale of operations within a company.

The designed system focuses on the importance of consolidating customer data spread across various departments and Enterprise Data Warehousing (EDW) so that all customer data can be accessed quickly and effectively. Machine learning is used to increase the revenue and scale of operations of a company in this system as a recommendation system with the decision tree method which produces an accuracy of 97.84% for training data and 95.38% for test data, to generate Business Insight.

Overall, this research explores the potential of artificial intelligence (AI) such as data mining strategies and data structures in the telecommunications industry. This data is also integrated into a secure unit, so that telecommunications service user data can be protected from data confidentiality problems. The results of data integration are also visualized in the form of distribution of user data using map charts, pie charts and bar charts. The resulting data visualization is in the form of code embedded on a PHP-based website using the CodeIgniter framework. This solution can help improve business strategy, predict customer behavior based on data visualization on the website and the recommendation system in it, so as to increase company sales in this competitive market.

Keywords: Business Insight, data confidentiality, Data Fabric, data integration, machine learning, system, website.