CHRIST
(DEEMED TO BE UNIVERSITY)
BANGALORE · INDIA

Notice for the PhD Viva-Voce Examination

Mr Sajanraj T D (Registration Number: 1970077), PhD scholar at the School of Engineering and Technology, CHRIST (Deemed to be University), Bangalore will defend his PhD thesis at the public viva-voce examination on Friday, 2 August 2024 at 10.00 am in the CDI Conference Room, III Floor, Block V, Bangalore Kengeri Campus, Bengaluru 560074.

Title of the Thesis	:	**Large Scale Transportation Data Analysis and Distributed Computational Pipeline for Optimal Metro Passenger Flow Prediction**
Discipline	:	**Computer Science and Engineering**
External Examiner (Outside Karnataka)	:	**Dr Ramesh V** Professor School of Computer Science and Engineering VIT - AP University, Inavolu Amaravati Andhra Pradesh - 522237
External Examiner (Within Karnataka)	:	**Dr Indiramma M** Professor and Head Department of AI & DS BMS College of Engineering Bull Temple Road, Basavanagudi Bengaluru - 560019 Karnataka
Supervisor	:	**Dr Raghavendra S** Associate Professor Department of Computer Science and Engineering School of Engineering and Technology CHRIST (Deemed to be University) Bengaluru – 560074 Karnataka

The members of the Research Advisory Committee of the Scholar, the faculty members of the Department and the School, interested experts and research scholars of all the branches of research are cordially invited to attend this open viva-voce examination.

Registrar

Place: Bengaluru
Date: 24 July 2024

ABSTRACT

Transportation has a significant impact on controlling traffic around a busy city. Among the transport systems, metro rails became the backbone of transport by operating above the traffic. For this reason, we have to take special consideration of the passenger flow in the transport system and, by understanding the needs, take timely actions for smooth running. Every metro system stores information about the flow of passengers in the form of transactions known as Automatic Fare Collection (AFC) data. For this research, AFC data is taken as the primary source of information to identify the passenger flow within the metro rail platform. Each metro system generates massive data throughout its running period and stores data for future reference. The information from the data source must be extracted and analyzed to provide forecasting. Considering the size of the data generated, the analytic platform has to process them in a distributed paradigm to handle "Big Data". Artificial Intelligence (AI) algorithms can derive information, insights, and patterns from this data. The patterns in the time series can be identified from the passenger flow data using exploratory data analysis. Exploratory data analysis is an essential step in data science for understanding the underlying properties of the raw data.

The research uses a data platform with a distributed computing and storage mechanism called the JP-DAP. The research leverages the above mentioned platform to extract passenger flow data from AFC Ticketing data. After the data engineering, the results of passenger flow information underwent further visualization and trend analysis. Based on the facts or patterns identified from the passenger flow information, a decision is taken for passenger forecasting. The initial study will reveal the characteristics of metro usage and practices within the system and finally derive a solution with a machine learning-based forecasting solution. The passenger flow forecasts based on the above patterns depend on factors like seasonality, trends, cyclicity, location, events, and random effects. By identifying the relative importance of these external factors, it can be integrated into the passenger flow forecast to improve the prediction accuracy of the system. This research identifies the station encoding as an essential factor that improves the forecasting performance. Thus, a station memorizing LSTM proposed improving the forecast performance with a minimum mean squared error rate of 0.0015. pattern recognition and passenger forecasts help to optimize the services aligned with the metro service. This gives a baseline for rescheduling the service to meet passenger demand, and it can benefit the metro in terms of cost, energy, and staffing deployment.

Keywords: Data Science, Distributed Computing, Passenger Flow, Short-Term Forecast

Publications:

1. **Sajanraj T. D.**, Mulerikkal, J. & S, Raghavendra Operational pattern forecast improvement with outlier detection in metro rail transport system. Multimed Tools Appl (2023), pp.1-17, DOI:10.1007/s11042-023-15637-x.
2. **Sajanraj, T. D.**, Mulerikkal, J., Raghavendra, S., Vinith, R., & Fabera, V. (2021). Passenger flow prediction from AFC data using station memorizing LSTM for metro rail systems. Neural Network World, 31(3), 173. DOI : 10.14311/NNW.2021.31.009.