ROBERT GORDON UNIVERSITY ABERDEEN

SCHOOL OF COMPUTING SCIENCE AND DIGITAL MEDIA

INFORMATICS INSTITUTE OF TECHNOLOGY

Student Name: D M D K Jayasena	Matriculation Number:
Supervisor: MR. DINESH ASANKA	Second Marker:
Project Title: Predicting raw material supplier delivery lead time variation in the apparel industry in Sri Lanka.	
MSc Business Analytics	Start Date: **30th September 2020**
	Submission Date: **30th April 2021**

CONSENT

I agree ☒

I do not agree ☐

That the University shall be entitled to use any results, materials or other outcomes arising from my project work for the purposes of non-commercial teaching and research, including collaboration.

DECLARATION

I confirm:

- **That the work contained in this document has been composed solely by myself and that I have not made use of any unauthorized assistance.**
- **That the work has not been accepted in any previous application for a degree.**
- **All sources of information have been specifically acknowledged and all verbatim extracts are distinguished by quotation marks.**

Student Signature:	Date Signed:

Abstract

In the apparel manufacturing industry, the lead time is the time it takes from the initial design phase to the final ready-made garment in the store, in other words, the total lead time is an important factor as the fashions are changing rapidly. On the other hand, to the apparel manufacturer with their make to order business strategy they don't have much time to act as the raw material ordering starts with the confirmed order from the customer. Having said that delivering the garments to the customer on the agreed date remains the same. Among other things like men, machine and methods availability of raw materials to perform the manufacturing is a key factor as even one raw material unavailability will hold the entire manufacturing of that particular style. As such lead times of the raw material supplier lead-times play a critical role in this industry. Even though there are multiple types of research conducted in lead-time prediction almost all of them were tried to predict the manufacturing lead time not the lead time of the supplier. On the other hand, there was no research conducted concerning the apparel manufacturing industry according to the literature. As such, there is a research gap in the supplier lead time prediction in the industry which leads to a perfect background for this research. Supplier lead time variation prediction in both aspects which are early receipt and delays are having an impact on the apparel manufacturer as the delays will cause the production to be delayed and idle time on the other hand early receipt will cause storage issues and the working capital management issues.

Real word data have been used in this research data masking have been used to maintain the confidentially of the data, exploratory data analysis have been carried out to understand the data extracted for the model building and null values and other company-specific filters were used to remove unnecessary observations. Extreme values were removed by calculating the percentiles where percentiles over 99% and below 1% were removed. Data preparation techniques such as feature engineering and feature selection were used to generate new features and remove features that are not important for the model building. Data transformation steps like log transformation of numerical variables were used to handle the skewness of the data. One hot encoding and target mean encoding techniques were used to convert the categorical variable to numerical variables where machine learning models can be used. In addition, feature selection and dimensionally reductions steps have been applied in this research, where features that have higher correlations (more than 0.7) have been removed from the feature space in order to handle any overfitting of the models.

Since the variable to a predicted is no of days delay or early receipt of raw materials this problem is a regression model in ML and 4 Models (Random Forest, XGBoost, Adaboost and Neural network) used. Hyperparameter tuning has been done manually on an iterative basis as well as automatically using the grid search technique in Python sklean. Based on the error measures (MSE, MAE and R2) random forest outperformed other models. In conclusion, ML can be used to predict the supplier lead time in the apparel industry in Sri Lanka. This research can be further improved by adding external features into the feature space and using the big data processing technique to train models with a large amount of data in the future.