A logo for college computing

Description automatically generated

**Assessment Cover Page**

**Declaration**

By submitting this assessment, I confirm that I have read the CCT policy on academic misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source.

I declare it to be my own work and that all material from third parties has been appropriately referenced.

I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution.

|  |  |
| --- | --- |
| *Student Full Name* Jason Liu Doyle |  |
| *Student Number* 2024200 |  |
| Module Title  Strategic Thinking |  |
| *Assessment Title*  Capstone Project proposal |  |
| *Assessment Due Date 29/03/2024* |  |
| *Date of Submission*  *29/03/2024* |  |

Contents

[***Optimizing Supply Chain Efficiency through Machine Learning: A Comprehensive Approach to Predictive Analytics, Customer Segmentation, and Demand Forecasting.*** 1](#_Toc162622781)

[Introduction 1](#_Toc162622782)

[Objectives 2](#_Toc162622783)

[Problem Domain 2](#_Toc162622784)

[Data sources & Ethical considerations 2](#_Toc162622785)

[Scope and methodology 3](#_Toc162622786)

[In depth analysis and timeline 3](#_Toc162622787)

[Tangible Results 4](#_Toc162622788)

[References 4](#_Toc162622789)

# ***Optimizing Supply Chain Efficiency through Machine Learning: A Comprehensive Approach to Predictive Analytics, Customer Segmentation, and Demand Forecasting.***

# Introduction

I plan to do a project on building a supply chain optimising model with machine learning, this will involve 3 components that will interlink and inform each other. I believe my project is important because it addresses critical inefficiencies in supply chain management.

My project is important because it addresses critical inefficiencies in supply chain management through the application of machine learning, offering solutions to reduce late deliveries, optimize inventory, and enhance customer satisfaction which can directly tie into a company’s financial optimisation (A. Agus 2013) in a sector that is fundamental to global commerce (Baldwin, 2012)

It’s relevance lies in how it addresses issues in the current need for businesses to adapt to the rapidly changing market dynamics and consumer expectations (Makris et al., 2019)

I think what makes this project interesting is exploring the relation of fairly pivotal aspects of an organisation:

* Customer Segmentation (Cooil et al., 2008)
* Predictive Analytics (Eckerson, 2007)
* Demand Forecasting (Armstrong et al.,2005)

and playing with how they interlink and how they’re improvement can improve an organisation.

I believe that this project will be able to be directly translated into real world application of business strategies in a broad range of sectors but a varying size of organisations. Its inclusion of data visualisation and actionable insight step at the end makes it highly accessible to people outside of the data analytics department.

# Objectives

The objective of my project will be the following;

* **Reduce Late Deliveries** through implementing some predictive analytics to accurately forecast potential delays in the supply chain, which will hopefully result in reducing the rate of late deliveries by a significant percentage.
* **Optimize Inventory Management** by using demand forecasting models to match inventory levels closely to real-time demand, which would reduce both overstock and stock shortages for an organisation.
* **Enhance Customer Segmentation** by apply machine learning algorithms for detailed customer segmentation, which would allow personalized marketing and customer service strategies to improve customer satisfaction and loyalty.
* **Explore the hypothesis** that integrating machine learning into an organisation’s supply chain operations would significantly improve operational efficiency, reduce costs, and enhance customer service, which would provide a competitive advantage in the marketplace.

# Problem Domain

My capstone proposal addresses supply chain inefficiencies in an organisation. It focuses on reducing occurences of late deliveries, optimising overall inventory management and improving customer segmentation. These inefficiencies impact a business’s operational costs, customer satisfaction, and overall, their revenue. Fixing these issues through optimising inefficient processes is vital for competitive advantage and keeping business sustainability in a globalised marketplace.

# Data sources & Ethical considerations

My dataset is called "DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS" and I found it on Kaggle (Shivp, 2020). This dataset, ideal for analysing inventory management, customer segmentation, and delivery times, has 53 columns and 180,520 rows. Access to the dataset is granted under a license allowing for sharing, adaptation, and commercial use, with the condition of proper attribution. Its open access makes sure the project abides by legal and ethical standards. I will focus on:

**Transparency** - I will document all data sources, methods and modifications in my project

**Data Anonymity** - Even though there isn’t any identifiable information currently in my chosen dataset I will make sure to anonymise any additional data that may be used if it becomes identifiable

**Adherence to Data Use Agreement** - I will respect the dataset's licensing terms, including the accurate attribution and following the outlined freedoms and restrictions, to make sure I adhere to the creators' intentions and legal requirements.

**Ethical Impact Assessment**: I will, throughout the project, keep an eye on the process to ensure no biases or inequalities affect the outcome.

# Scope and methodology

The purpose of my project is to use machine learning to optimise a supply chain, specifically by focusing on inventory management, customer segmentation and delivery times.

**I will include:**

* Definition of supply chain inefficiencies and project objectives.
* Analysis of dataset focusing on
  + Late deliveries – (Predictive analytics)
  + Customer Segmentation – (Improve targeted marketing)
  + Inventory levels – (Demand forecasting)
* Evaluation of findings
* Conclusion with actionable insights and supporting visualisations

**I will exclude:**

* Macroeconomic factors
* Competitor analysis
* Geographical scope
* Industry specific factors

**My boundaries are:**

* Timeframe of only 2 semesters
* Lack of context for model in a real life environment
* Not able to access real-time data.

# In depth analysis and timeline

**Semester 1**

* Data Collection and Pre-processing – 5 weeks
  + Preparing the data in Jupyter notebook through data cleaning, normalisation and transformation.
* Exploratory Data Analysis (EDA) – 6 weeks
  + Where I get to understand the data characteristics, gathering my findings with python to conduct corelation analysis and statistical summaries.
* Model Development - 8 weeks
  + Creating the machine learning models. Initial model training and parameter tuning using algorithms suitable for regression (for forecasting) and classification (for predictive analytics and segmentation).

**Semester 2**

* Model Development (Continued) – 5 weeks
  + I am leaving ample time for this part as I may struggle the most with it and it may take me the longest.
* Model Evaluation and Refinement - 6 weeks
  + Assessing and improving model performance through cross validation and testing.
* Analysis and Interpretation – 6 weeks
  + Deriving insights from my model results and translating them into actionable insights for management to use.
* Finalization and Presentation – 3 weeks
  + Compiling and presenting the findings and making it accessible for all departments (non-data analysts)

# Tangible Results

By the end of semester two I wish to have:

1. **Analysis and Models**: A detailed report which will summarize:
   * Predictive analytics for late deliveries.
   * Demand forecasting for inventory optimization.
   * Customer segmentation for marketing strategies.
   * Model effectiveness in supply chain improvements.
2. **Strategic Recommendations**: Which will hopefully include
   * Strategies to reduce late deliveries.
   * Inventory and marketing recommendations.
   * Implementation framework for real-world application by management

# GitHub Link

https://github.com/CCT-Dublin/capstone-project-feb-2024-pt-s17characters.git

# References

A. Agus, “The Importance of Supply Chain Management on Financial Optimization”, *Jurnal Teknik Industri: Jurnal Keilmuan dan Aplikasi Teknik Industri*, vol. 15, no. 2, pp. 77-84, Dec. 2013.

Armstrong, J.S. and Green, K.C., 2005. *Demand forecasting: evidence-based methods* (No. 24/05). Monash University, Department of Econometrics and Business Statistics.

Baldwin, R.E. (2012) *Global Supply Chains: Why they emerged, why they matter, and where they are going*, *SSRN*. Available at: https://ssrn.com/abstract=2153484 (Accessed: 19 March 2024).

Cooil, B., Aksoy, L. and Keiningham, T.L. (2008) ‘Approaches to customer segmentation’, *Journal of Relationship Marketing*, 6(3–4), pp. 9–10. doi:10.1300/j366v06n03\_02.

Eckerson, W.W. (2007) *PREDICTIVE ANALYTICS Extending the Value of Your Data Warehousing Investment*. Available at: http://download.101com.com/pub/tdwi/files/pa\_report\_q107\_f.pdf (Accessed: March 2024).

Makris, D., Hansen, Z.N. and Khan, O. (2019a) ‘Adapting to supply chain 4.0: An explorative study of multinational companies’, *Supply Chain Forum: An International Journal*, 20(2), pp. 116–131. doi:10.1080/16258312.2019.1577114.

Shivp (2020) *Dataco Smart Supply Chain for Big Data Analysis*, *Kaggle*. Available at: https://www.kaggle.com/datasets/shivkp/customer-behaviour (Accessed: 29 March 2024).