



AI-Powered Fashion Trend Prediction: A Conceptual Framework and Comparative Analysis Using Social Media and Historical Data

K.T.S.P. De Noyel
(MS24019804)

A THESIS
SUBMITTED TO
SRI LANKA INSTITUTE OF INFORMATION TECHNOLOGY
IN PARTIAL FULFILMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF SCIENCE IN INFORMATION MANAGEMENT

December 2025

I certify that I have read this thesis and that in my opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Ms. Anjalie Gamage
Senior Lecturer, Faculty of Computing, SLIIT

Approved for MSc. Research Project:



MSc in IM Program Co-Ordinator, SLIIT

Approved for MSc:

Head of Graduate Studies, FoC, SLIIT

DECLARATION

This is to certify that the work is entirely my own and not of any other person, unless explicitly acknowledged (including citation of published and unpublished sources). The work has not previously been submitted in any form to the Sri Lanka Institute of Information Technology or to any other institution for assessment for any other purpose.

A handwritten signature in black ink, appearing to read 'A. Hashem', with a stylized flourish extending from the end.

Sign:

Date: 02.11.2025

ABSTRACT

Fashion trend prediction is a complex process influenced by various factors have a role to play in. With emphasis on data observation and collection within an industry, interest in predicting trends using artificial intelligence (AI) has intensified over time. This paper outlines a theoretical framework of an AI-driven fashion trend prediction system that combines past sales data and social media posts. The model aims to use natural language processing (NLP) to explore Instagram, Facebook, fashion forums, and other fashion-related social media platforms to process information to capture the real-time opinion of the masses and identify trends. [1] At the same time, historical sales performance data, including products sold during seasons, subsequent years, purchasing behavior over time, and other consumer and market trends, assist in predicting demand in the marketplace. [2]

The suggested framework describes how AI algorithms, especially machine learning and deep learning models can be used to these two complementary data sources to create actionable insights. When these datasets are combined, fashion trends are forecasted with more accuracy. This will assist brands and retailers in making proactive decisions regarding product and marketing strategies alongside inventory control. In particular, the paper underscores the conceptual architecture of the AI system, explaining the fundamental aspects of forecasting trends, consolidating data, and computation through an algorithm. This forms the basis of other studies that seek to apply AI Systems in forecasting the fashion industry trends while at the same time waiting to see the deployment of such systems. In addition, the challenges alongside ethical implications as well as opportunities for further development within this context strive to reduce the existing gap between fashion business practices and data science are provided.

ACKNOWLEDGEMENT

I would like to say a profound thanks to everyone who helped me and showed me the way during the period of this research. To begin with, I would like to thank my supervisor, Ms. Anjalie Gamage, whole-heartedly because of his priceless mentorships and feedback. I would also like to express my gratitude to the Sri Lanka Institute of Information Technology (SLIIT) which has offered me a thought provoking and resourceful environment in which I am able to complete this study.

A special thanks to Dr. Prasanna S. Haddela for his guidance on this research project. I also extend my heartfelt gratitude to the panel of presentation Prof. Anuradha Jayakody, Lecturer. Mr. Senior Lecturer Kanishka Yapa. Ms. Senior Lecturer and Suranjini Silva. Ms. Narmada Gamage, and the other lecturers who treated their time, positive comments and valuable suggestions as valuable assets to this work, which made great contribution in enhancing the quality of this work.

I also owe great debt to my colleagues, friends, family and peers who provided support and encouraged me in the process of data collection and analysis, as well as helped me in many aspects. Their contributions are too numerous to mention separately, and they have been invaluable.

Lastly, my parents have contributed to my great academic success as I could not have been as successful without their unwavering support, patience, and faith in me. They are my source of strength and encouragement since their love, sacrifices, and always-encouraging attitude have been my greatest strength and motivation.

Table of Contents

DECLARATION	ii
ABSTRACT	iii
ACKNOWLEDGEMENT	iv
Table of Contents	v
List of Figures	ix
List of Tables	xi
Chapter 1 Introduction	1
1.1 Background	1
1.2 Research Problem	2
1.2.1 Research Problem 1	4
1.2.2 Research Problem 2	4
1.2.3 Research Problem 3	4
1.2.4 Research Problem 4	4
1.3 Research Objectives	6
1.4 Conceptual Framework	7
1.5 Research Hypotheses	9
1.5.1 Hypotheses	9
1.6 Research Questions	11
1.6.1 Research Questions 01	11
1.6.2 Research Questions 02	11
2.6.3 Research Questions 03	11
1.6.4 Research Questions 04	12
1.7 Significance of the Study	12
1.8 Expected Contributions	13
1.8.1 Future Research Areas	14
1.9 Limitations	15

Chapter 2 Literature Review.....	17
2.1 Review of Existing Literature	17
2.2 Gap Identification.....	19
Chapter 3 Research Methodology	22
3.1 Research Design.....	22
3.1.1 Independent variables (IV)	23
3.1.2 Moderating and Mediating Variables (M).....	24
3.1.3 Control Variables (C)	25
3.2 Conceptual framework	26
3.2.1 Conceptual and Statistical Framework of Demographic Analysis	27
3.2.2 Conceptual and Statistical Framework of Social Media Data Analysis.....	29
3.2.3 Conceptual and Statistical Framework of AI Acceptance Personalization	30
3.2.4 Conceptual and Statistical Framework of ARIMAX	32
3.2 Sampling Strategy	33
3.2.1 Target Population	33
3.2.2 Sample Size	34
3.3 Survey Participants & Inclusion & Exclusion Criteria	35
3.4 Data Collection Methods.....	36
3.4.1 Surveys	36
3.4.2 Social media data extraction.....	36
3.4.3 Data Analysis Techniques	38
Chapter 4 Data Analysis	40
4.1 Data Analysis Methods	40
4.2 Data Preparation and Cleaning	40
4.2.1 Survey Data	40
4.2.2 Social Media Data	41
4.2.3 Sales Data	41
4.3 Analytical Techniques.....	42
4.3.1 Statistical Description.....	42
4.3.2 Reliability & Validity Tests.....	42

4.3.3 Social Media Data Sentiment Analysis	42
4.3.4 Inferential & Predictive Modeling.....	42
4.3.5 Model Validation.....	43
4.4 Demographic Analysis	43
4.4.1 Sample Characteristics	45
4.4.3 Age as a Predictor of Purchase Intensity	54
4.5 Social Media Data and Sentiment Analysis	63
4.5.1 Sentiment Approach	63
4.5.3 Time-Series & Decomposition	64
4.5.3 Engagement Analysis - Cargo vs Skinny Jeans	66
4.5.4 Engagement Analysis - Vintage Vs Chunky Sneakers.....	74
4.5.4 Hashtag Analysis	80
4.5.5 Hashtag Velocity as a Predictor for Purchase Intensity	82
4.6 AI Acceptance Analysis.....	87
4.6.1 Cronbach's Alpha Reliability Analysis Description and Insight	87
4.6.2 Standard Heat Map Analysis Description and Insight	90
4.6.2 Statistical model AI Acceptance Trust and Trend Adoption.....	92
4.6.3 Relationship Between Personalization Accuracy and Purchase Intensity.....	95
4.7 Integrated Framework.....	97
4.7.1 Aligning ARIMA with Sentiment analysis	97
4.8 Modal Validation.....	105
4.8.1 Model Validation Engagement Analysis with Actual Market	105
4.8.2 ARIMAX vs ARIMA Model Performance Analysis	106
Chapter 5 Conclusion	112
5.1 Correlation Analysis and Mediation Insights	114
5.2 Engagement Patterns	114
5.3 AI Trust, Acceptance, and Trend Adoption	115
5.4 Managerial and Theoretical Implications	116
5.5 Ethical Concerns.....	117
5.6 Expert Review and Amendments.....	118

References	120
Appendix.....	123
Appendix 1: Mapping of Variables, Hypotheses, Research Questions, and Survey Items	123
Appendix 2: Survey Questionnaire.....	124
Appendix 3: JSON Input for Apify Data Scraper.....	133
Appendix 4: Expert Review and Recommendation	136
Expert Details – Reviewer 1	136
Expert Details – Reviewer 2	141

List of Figures

Figure 1.1 Variable Hypothesis Mapping Diagram.....	8
Figure 3.1 Conceptual Framework	26
Figure 3.2 Conceptual Modal Section 1 Demographic Data.....	27
Figure 3.3 Statistical Modal Section 1 Demographic Data	28
Figure 3.4 Conceptual Modal Section 2 Social Media Data	29
Figure 3.5 Statistical Modal Section 2 Social Media Data.....	29
Figure 3.6 Conceptual Modal Section 3 AI Acceptance Personalization.....	30
Figure 3.7 Statistical Modal AI Acceptance & Trend Adoption	31
Figure 3.8 Statistical Model ARIMAX	32
Figure 3.9 A Few Examples of Social Media Data Extraction from Apify	37
Figure 3.10 A Few Examples of Social Media Hashtag Scraper from Apify	38
Figure 4.1 Demographic characteristics	44
Figure 4.2 Income vs Social Media Platform Usage	45
Figure 4.3 Boxplot: Purchase behaviors by Income Band	45
Figure 4.4 Mediation Boxplots Purchase Intensity by Gender & Income.....	48
Figure 4.5 Mediation Correlation Heatmap.....	49
Figure 4.6 Mediation Scatter Plots	51
Figure 4.7 Purchase Behavior by Income Boxplot.....	53
Figure 4.8 Purchase Intensity Distribution by Age	56
Figure 4.9 Age Purchase Distributions and Residuals.....	57
Figure 4.10 Mean Purchase Intensity by Age.....	59
Figure 4.11 Age vs Purchase Intensity	60
Figure 4.12 Instagram sentiment distributed were classified into three categories	63
Figure 4.13 Time Series Analysis Illustrating the Trend within the Sustainable Fashion Theme	65
Figure 4.14 Time Series Decomposition Sustainability Fashion Theme	65
Figure 4.15 Statistical Hypothesis Testing	68
Figure 4.16 Engagement Distribution Cargo vs Skinny Boxplots	69
Figure 4.17 Mean Engagement Visualization Cargo vs Skinny.....	70
Figure 4.18 Cargo vs Skinny Distributions	71
Figure 4.19 Scatter Plot Relationship Between Likes and Comments Cargo vs Skinny.....	73
Figure 4.20 Brand Visibility & Audience Interaction: Chunky vs Vintage Sneakers Box Plots	76
Figure 4.21 Mean Engagement Metrics: Chunky vs Vintage Sneakers	77
Figure 4.22 Engagement Distribution Analysis: Chunky vs Vintage Histogram Violin Plot	78
Figure 4.23 Engagement Distribution Chunky vs Vintage Scatter Plot	79
Figure 4.24 Causality network of denim-themed statistics	82
Figure 4.25 Boxplots Diagram Hashtag Velocity.....	84
Figure 4.26 Hashtag Velocity Correlation.....	84
Figure 4.27 Hashtag Velocity Prediction.....	85
Figure 4.28 Cronbach's Alpha values	87
Figure 4.29 Spearman Correlation Heatmap for Mediation Variables.....	92
Figure 4.30 3D Visualization Personalization Trust Purchase	96
Figure 4.31 The Computed Pseudo-demand Variable (D_t)	99
Figure 4.32 ARIMAX Modal Components	100

Figure 4.33 ARIMAX Actual vs Predicted Sales with Sentiment.....	101
Figure 4.34 ARIMAX Model Actual vs Predicted.....	102
Figure 4.35 ARIMAX Residual Distribution	104
Figure 4.36 Market Dynamics Chunky vs. Vintage Sneaker Performance.....	105
Figure 4.37 ARIMA ARIMAX Comparison.....	108
Figure 4.38 ARIMA ARIMAX Model Performance	110
Figure A.1 JSON Input for Cargo Denim Hashtags Scraper.....	133
Figure A.2 JSON Input for Skinny Jeans Hashtags Scraper	134
Figure A.3 JSON Input for Vintage Sneakers Hashtags Scraper	134
Figure A.4 JSON Input for Chunky Sneakers Hashtags Scraper	135
Figure A.5 Expert Review Email – Evaluation and Recommendations.....	140
Figure A.6 Expert Review Email – Evaluation and Recommendations.....	142

List of Tables

Table 4.1 Purchase Intensity Segments Analysis	47
Table 4.2 Purchase Intensity by Age Group	55
Table 4.3 One-tailed Two-sample t-test Results.....	66
Table 4.4 Linear Regression Results	67
Table 4.5 Table 0.2 Linear Regression Results	67
Table 4.6 Descriptive Statistics Across Two Sneaker Categories	75
Table 4.7 Normality and Inferential Test Results.....	75
Table 4.8 Cohen’s d Effect	76
Table 4.9 Summary table showing the sentiment polarity	81
Table 4.10 Statistical Hypothesis Results.....	83
Table 4.11 Type differences (Cargo vs. Skinny).....	83
Table 4.12 Measurement Scales and Constructs for Cronbach’s Alpha	88
Table 4.13 Description of the Chart	89
Table 4.14 Correlation matrices in statistics.....	90
Table 4.15 The numerical values used for the analysis are based on the ordinal scales	94
Table 4.16 The total effect (c) is the sum of the direct and indirect effects	94
Table 4.17 Metrics that reflect consumer interest	98
Table 4.18 Weighted Engagement acts as a demand signal.	98
Table 4.19 responsiveness weights.....	99
Table 4.20 Predictive alignment and correlation.....	100
Table 4.21 ARIMAX vs ARIMA Model Performance Analysis	106
Table A.1 Mapping of Variables, Hypotheses, Research Questions, and Survey Items.....	123