

Strategic Thinking – Final Report

Student: Renata Oliveira Souza

SECTION 1 — INTRODUCTION

The fashion and activewear industry has undergone rapid transformation in recent years, driven by digitalisation, social media engagement, and a growing consumer interest in health, fitness and body-enhancing products. Anti-cellulite leggings, in particular, have become a high-demand product category due to their aesthetic appeal, compression benefits, and strong presence in influencer-driven promotional campaigns. In this context, strategic decision-making increasingly depends on data availability, analytical capability and technological readiness. Businesses that effectively integrate information systems, analytics and market intelligence gain a clear competitive advantage.

This report forms part of the Strategic Thinking module and builds upon the capstone project developed in semester one. Whereas earlier assessments focused primarily on data preparation and machine learning implementation, the present report expands the perspective to critically evaluate the project through a strategic, technological and organisational lens. The goal is not only to analyse customer behaviour but also to understand how IT infrastructure, analytical tools, and emerging technologies can support long-term competitiveness in the activewear sector.

The project uses the Women's E-Commerce Clothing Reviews dataset from Kaggle, complemented by additional synthetic variables—Price, Customer Country and Sales Channel—modelled to reflect realistic patterns observed in online fashion retail. The artefact includes detailed exploratory data analysis (EDA), the development and tuning of three machine learning models, and model interpretability to identify the drivers behind customer recommendations. However, in the context of this module, the technical artefact serves as evidence for a broader strategic assessment.

The purpose of this report is therefore fourfold: (i) to critically evaluate the relationship between IT infrastructure and organisational competitive advantage; (ii) to analyse the technology decisions made in the project; (iii) to examine the business environment using strategic analysis tools; and (iv) to assess the ethical, societal and organisational implications of using customer data and predictive models in a consumer-facing domain.

SECTION 2 — BUSINESS AND MARKET CONTEXT

The business scenario examined in this project is that of a digital-first fashion retailer specialising in anti-cellulite leggings. These products combine style, body support and smoothing technology designed to reduce the appearance of cellulite during exercise. Their appeal is amplified by the increasing visibility of fitness, wellness and athletic aesthetics promoted through Instagram, TikTok and YouTube. Consumers in this segment often demonstrate high engagement with social media, respond well to influencer endorsements and rely extensively on product reviews when making purchasing decisions.

The brand's value proposition is anchored in three elements: aesthetic enhancement, functional performance and digital convenience. The competitive environment is characterised by large sports brands, specialised fitness retailers and aggressive low-cost competitors. Given this context, predicting customer recommendations provides strategic value.

SECTION 3 — IT INFRASTRUCTURE AND COMPETITIVE ADVANTAGE

Information technology infrastructure enables modern organisations to collect, store, process and analyse data at scale. The infrastructure used in this project includes data sources, analytical processing tools (Python, pandas, scikit-learn), machine learning capabilities, and GitHub for version control. These technologies support fast experimentation, scalability, transparency and competitive advantage by transforming data into business insights.

SECTION 4 — TECHNOLOGY CHOICES: OPEN SOURCE VS PROPRIETARY

The project relies on open-source tools such as Python, Jupyter, pandas, NumPy, scikit-learn and GitHub. Advantages include cost reduction, flexibility, transparency, rapid innovation and global community support. Proprietary alternatives like SAS, SPSS, Tableau and Adobe Analytics offer strong interfaces but impose licensing costs, vendor lock-in and reduced customisation. For a digital-first fashion brand, open-source technologies offer superior strategic alignment.

SECTION 5 — STRATEGIC BUSINESS ANALYSIS (PESTEL + SWOT)

The PESTEL analysis reveals that political regulations, economic conditions, social fitness trends, technological innovations, environmental sustainability pressures and legal data protection requirements shape the activewear market. The SWOT analysis identifies strengths (product differentiation, digital distribution), weaknesses (competition, dependency on social media), opportunities (international expansion, influencer marketing) and threats (fast-fashion competitors, algorithmic changes, economic downturns).

SECTION 6 — METHODOLOGY

The methodological approach includes dataset enrichment (price, country, sales channel), data cleaning, feature engineering (review_length), EDA, machine learning (Logistic Regression, Random Forest, Gradient Boosting), hyperparameter tuning and cross-validation. Evaluation metrics included accuracy, precision, recall, F1 and AUC. The methodology ensures a structured and reproducible process aligned with business goals.

SECTION 7 — MACHINE LEARNING RESULTS AND VALIDATION

Logistic Regression served as a baseline. Random Forest improved predictive performance significantly. Gradient Boosting achieved the strongest results across all metrics, especially F1-score and AUC. Model comparison shows that Gradient Boosting is best suited for consumer behaviour prediction. Cross-validation confirmed model stability.

SECTION 8 — MODEL INTERPRETABILITY

Key predictors include rating, positive feedback count, review length, price, country and sales channel. Interpretability ensures fairness, supports ethical deployment, enhances trust and enables marketing teams to make informed decisions.

SECTION 9 — EMERGING TECHNOLOGIES

Emerging technologies relevant to the organisation include AI personalisation, social commerce, smart fabrics, NLP for review analysis, AR virtual try-on tools and cloud computing. These

technologies increase engagement, improve customer experience and support long-term competitiveness.

SECTION 10 — ETHICS, FAIRNESS AND TRANSPARENCY

Ethical considerations include responsible data usage, GDPR compliance, bias mitigation, interpretability and societal impact. The project applies synthetic variables to avoid privacy risks and monitors fairness across demographic dimensions.

SECTION 11 — PROJECT CHALLENGES AND MANAGEMENT

Challenges included dataset limitations, model performance variability, and translating modelling output into strategic recommendations. Project management followed a phased structure: planning, preparation, analysis, modelling, validation, reporting and quality assurance.

SECTION 12 — CONCLUSION

This project demonstrates how strategic thinking, analytics and ethical AI can support organisational growth in the activewear sector. Predictive modelling provides actionable insights for marketing campaigns, customer segmentation and product strategy. Open-source tools offer cost-effective infrastructure, while emerging technologies create new opportunities for innovation.

REFERENCES

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32.

Friedman, J.H. (2001). Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*, 29(5), 1189–1232.

Hastie, T., Tibshirani, R., Friedman, J. (2009). *The Elements of Statistical Learning*. Springer.

Kaggle (n.d.). Women's E-Commerce Clothing Reviews. Available at: <https://www.kaggle.com/>

Pedregosa, F. et al. (2011). *Scikit-learn: Machine Learning in Python*.

OpenAI (2025). ChatGPT (GPT-5.1). Used for clarification of code and methodology.