Classification of Brand Perception Using Random Forest: Brand Preference, Brand Loyalty, and Brand Trust

Muhammed İkbal Yılmaz, Özge Yıldız, and Şahika Gökmen

Ankara Hacı Bayram Veli University, Department of Econometrics

15-17 October, 2024

Contact: myucanlar@gmail.com

Summary

Introduction

Brand Preference (BP), Brand Loyalty (BL), and Brand Trust (BT);

- 'Brand Asset Management: Profitable Growth Through Your Brands' by Davis (1964), second edition.
- Personal values are important for BP as they fulfill individual needs and influence '**perceived quality**' (Ebrahim, 2013; Yang et al., 2015).
- Customer loyalty involves a repeated purchasing process (Gumus, 2016) and is influenced by ease of return and '**recognizability**' (Chen, 2021; Aaker, 1996).

Yang, et al., (2015):*Sustainable Brand Reputation: Evaluation of IPhone Customer Reviews with Machine Learning and Sentiment Analysis*.
Saylı, et al., (2016): *Brand Loyalty Analysis System Using K-Means Algorithm*.
Pamuksuz and Yun, (2021): *A Brand-new Look At You: Predicting Brand Personality In Social Media Networks With Machine Learning*.
Dong, (2023): *Application of User Preference Mining Algorithms Based on Data Mining and Social Behaviour in Brand Building*.

Methodology: Random Forest Classification

Machine Learning (ML);

- Turing (1950); McCarthy et al., (1955)

Random Forest Classification (Breiman, et al., 1984; Quinlan, 1986; Breiman, 1996; Breiman, 2001)

Step 1. Randomly sample data and independent variables to construct M decision trees using different subsets of the training dataset D_n.

Step 2. Each tree $h(x_i; \theta_j, D_n^{(j)})$ predicts the class of the independent variable x_i, based on the splitting criteria used for that specific tree.

Step 3. The final class prediction is determined by majority vote among the M trees:

$$\hat{y}(x_i) = \arg\max_{c \in C} \left\{ \sum_{j=1}^{M} 1 \left[h(x_i; \theta_j, D_n^{(j)}) = c \right] \right\}$$

Dataset and Implementation

Characteristics of the Dataset

Table 1 – Cronbach's Alpha.

	Alpha (α)
Overall	**0.8399**

Table 2 – Dataset Overview.

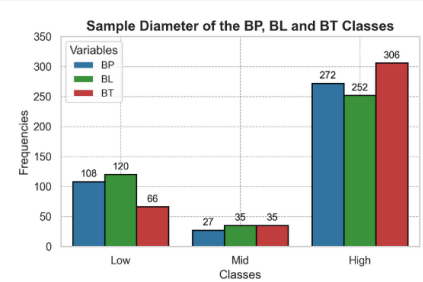
	Male	**Female**	n_i
Econometrics	65	99	164
Economics	75	108	243
		$\sum_{i=1}^{k} n_i$	**407**

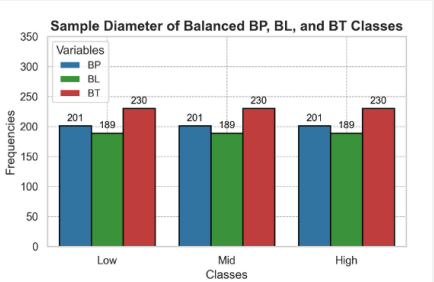
Table 3 – Variables and Classes.

Variables	**Classes**
Dependent	
BP	Low, Mid, High
BL	Low, Mid, High
BT	Low, Mid, High
Independent	
Age, Gender, Department, Year, BandR, Place, Income, Socialmedia, Ishopping, BrandMonitoring, Price, Quality, Timeofuse, ProductionDate, Seasonality, Material, Discount, Fashion, Design, Psychologicalimpulses, Country, Recognizability	

Dataset and Implementation

Original Dataset and Balanced Dataset





- The generated data originated from the original training set itself, meaning that no external information was introduced to the model.

Dataset and Implementation

Confusion Matrices and Performance Metrics;

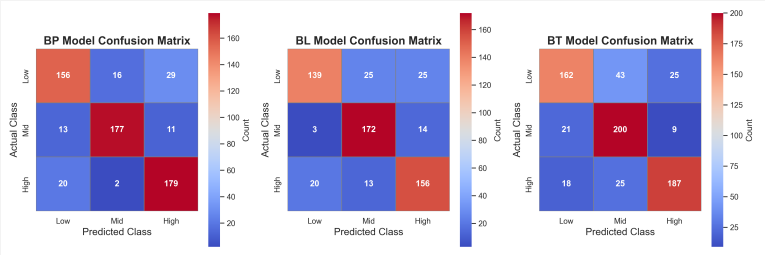
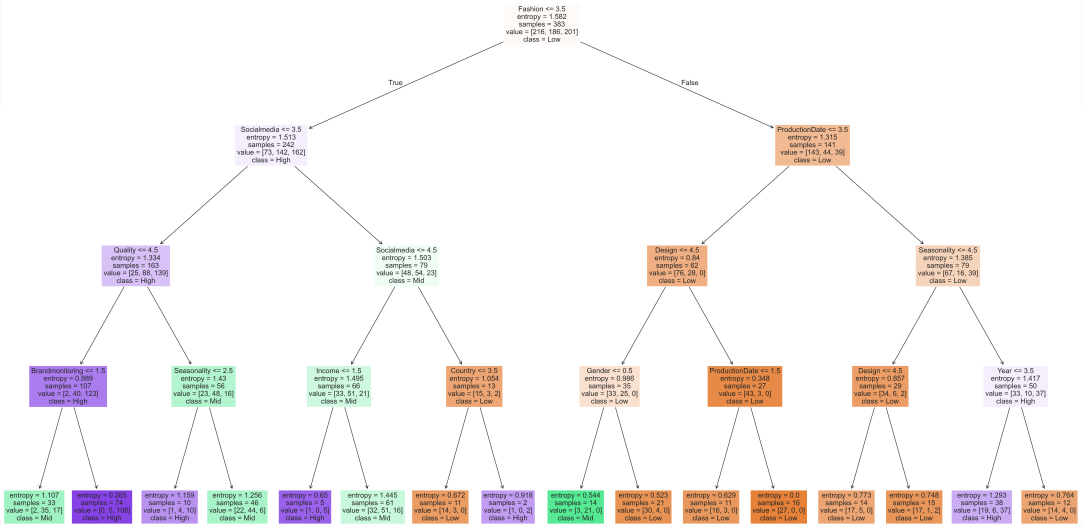


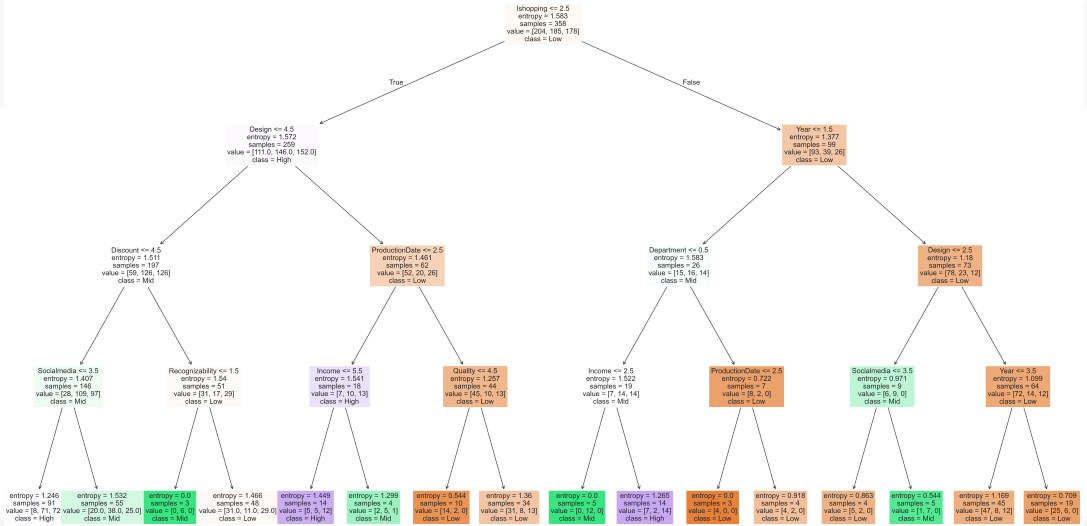
Table 4 – Model Performance Metrics.

	BP Model					BL Model					BT Model			
	Classes	Precision	Recall	F1		Classes	Precision	Recall	F1		Classes	Precision	Recall	F1
y_{BP}	Low	0.82	0.89	0.85	y_{BL}	Low	0.80	0.83	0.81	y_{BT}	Low	0.85	0.81	0.83
	Mid	0.83	0.78	0.80		Mid	0.86	0.74	0.79		Mid	0.81	0.70	0.75
	High	0.91	0.88	0.89		High	0.82	0.91	0.86		High	0.75	0.87	0.80
Accuracy: 0.85					**Accuracy:** 0.82					**Accuracy:** 0.80				

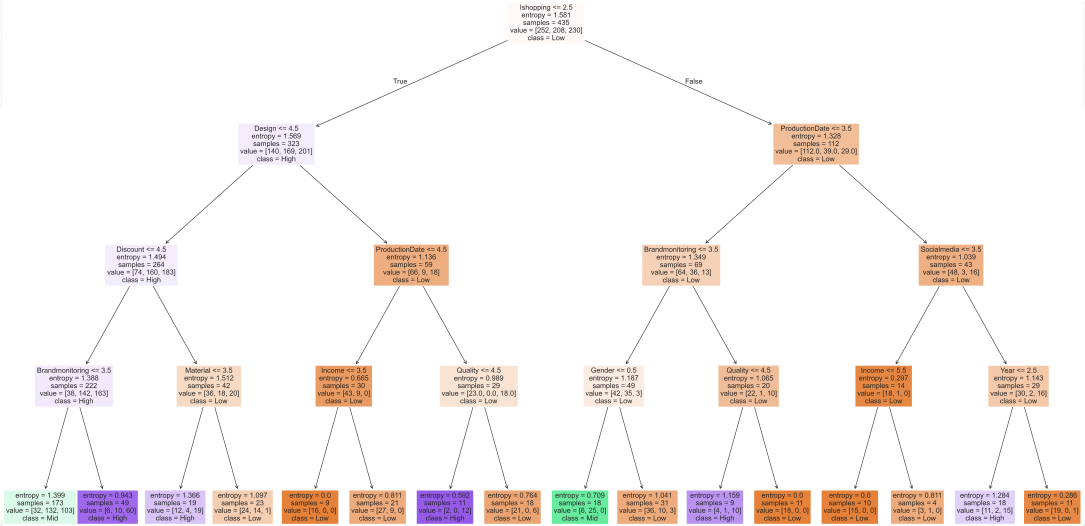
Dataset and Implementation: BP Model Decision Tree



Dataset and Implementation: BL Model Decision Tree

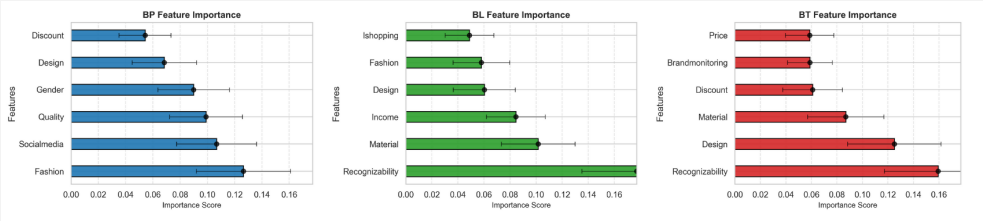


Dataset and Implementation: BT Model Decision Tree



Dataset and Implementation

Feature Importance Plots;



- It is confirmed that BP is influenced by many factors (Ebrahim, 2013).
- Recognizability drives '**situational**' loyalty rather than preference (Oliver, 2010; Yang, et al., 2015; Chen, 2021).
- Social Media offers an effective marketing and perception enhancement opportunity (Pamuksuz, et al., 2021).
- Technological innovation and knowledge-based service offerings, strategic advantages, (Khan and Rahman, 2016),
- Designs that combine functionality and aesthetics instill confidence and increase '**proactive**' loyalty (Oliver, 2010).

Results and Discussion

- The findings support previous research and contribute to the dynamics of the field,
- Online shopping platforms, Social Media, Age, Gender, Product Features,
- Artificial Intelligence Resources,
- Cost-Sensitive Learning, ADASYN, Tomek Links approaches can also be adopted.
- Lack of coefficient interpretation,
- A change of algorithm is of course also advisable,

References

Aaker, D. A. (1991). *Managing Brand Equity.* New York: The Free Press.
Aaker, D. A. (1996). *Building Strong Brands.* New York: Free Press.
Bastos, W., & Levy, S. J. (2012). *A History of The Concept of Branding: Practice and Theory.* Journal of Historical Research in Marketing, 4(3), 347–368.
Breiman, L. (1996). *Bagging Predictors.* Machine Learning, 24, 123-140.
Breiman, L. (2001). *Random Forests*. Machine Learning, 45(1), 5-32.
Breiman, L., Friedman, J., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees* (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315139470.
Chaudhary, A., Kolhe, S., & Kamal, R. (2016). *An Improved Random Forest Classifier for Multi-class Classification*. Information Processing in Agriculture, 3(4), 215-222.
Chawla, N. V., De Silva, V., & Van der Laan, M. (2004). *A Statistical Framework for Balancing Data Sets.* Proceedings of the 2004 American Association for Artificial Intelligence.
Chen, Y. (2021). *A Social Media Mining and Ensemble Learning Model: Application to Luxury and Fast Fashion Brands*. Information, 12(4), 149.
Davis, S. M. (1964). Brand Asset Management: Profitable Growth Through Your Brands (2nd ed.). San Francisco: Jossey-Bass.
Dong, Y. (2023). *Application of User Preference Mining Algorithms Based on Data Mining and Social Behaviour in Brand Building*.
Ebrahim, R. S. (2013). *A Study of Brand Preference: An Experiential View* (Doctoral dissertation, Brunel University Brunel Business School PhD Theses).
Gumus, I. (2016). *Brand Gender, Brand Personality and Brand Loyalty Relationship.* Communication in Mathematical Modeling and Applications, 1(2), 8-41.
Khan, I., & Rahman, Z. (2016). *E-tail Brand Experience's Influence On E-Brand Trust and E-Brand Loyalty: The Moderating Role of Gender*. International Journal of Retail & Distribution Management, 44(6), 588-606.
McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). *A Proposal for The Dartmouth Summer Research Project On Artificial Intelligence*, August 31, 1955. AI Magazine, 27(4), 12.
Oliver, R. L. (2010). *Consumer Brand Loyalty.* In Wiley International Encyclopedia of Marketing. Chichester, West Sussex, UK: Wiley-Blackwell, 432-436.
Oshiro, T. M., Perez, P. S., & Baranauskas, J. A. (2012). *How Many Trees In a Random Forest?* In Machine Learning and Data Mining in Pattern Recognition: 8th International Conference, MLDM 2012, Berlin, Germany, July 13-20, 2012. Proceedings 8. Springer Berlin Heidelberg.
Pears, R., Finlay, J., & Connor, A. M. (2014). *Synthetic Minority Over-sampling Technique (SMOTE) for Predicting Software Build Outcomes.* arXiv preprint arXiv:1407.2330.
Quinlan, J. R. (1986). *Induction of Decision Trees.* Machine Learning, 1, 81-106.
Samuel, A. L. (1959). *Some Studies In Machine Learning Using The Game of Checkers.* IBM Journal of Research and Development, 3(3), 210-229.
Turing, A. M. (1950). *Computing Machinery and Intelligence.* Mind, 59(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433.
Wang, J., Yao, Y., Zhou, H., Leng, M., & Chen, X. (2013, December). *A New Over-sampling Technique Based On SVM for Imbalanced Diseases Data.* In Proceedings 2013 International Conference on Mechatronic Sciences, Electric Engineering and Computer (MEC) (pp. 1224-1228). IEEE.
Yang, C., Pan, S., Mahmud, J., Yang, H., & Srinivasan, P. (2015). *Using Personal Traits for Brand Preference Prediction*. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (pp. 86-96).