MODEL PREDICTION OF THE NEXT RUNWAY MODEL USING DECISION TREE AND RANDOM FOREST

(CASE STUDY: BIG FOUR FASHION WEEK)

Dwita Adilah[1], Andry Alamsyah[2]

[1] S1 International ICT Business, School of Economic and Business, Telkom University
[1] School of Economic and Business, Telkom University
[1]dwitaadilah@student.telkomuniversity.ac.id, [2]andrya@telkomuniversity.ac.id

Abstract

Fashion industry contributes 4% of global market share with USD 385.7 billion-dollar market value worldwide. Fashion Week, the most prominent event in fashion industry, is a combination between art and commerce. Fashion models is the talent that plays a big role during the fashion week, in which their aesthetic is highly associated in presenting designers' work. However, their appearance in social network has built a deeper relationship to the industry. Choosing the talent out hundreds of faces is a challenge for casting director. The fast movement of fashion business has never evolved closely to digital utilization. This study aims to imitate the wisdom of traditional talent scouting process into an automation model based on machine learning practice by implementing Decision Tree, and Random Forest. In the era of fashion 4.0, shifting the traditional system into an automation model leads to revolution of fashion production, in this case fashion show production. Our framework is able to imitate talent scouting process to select the upcoming fashion model models appear on 2019 Fall/Winter Fashion Week with accuracy 74.57% using Random Forest.

Keywords: ***Fashion Industry, Machine Learning, Prediction, Decision Tree, Random Forest.***

1. Introduction

Fashion is a multi-billion-dollar industry with social and economic implication worldwide [1]. Fashion industry values three trillion US dollars with 4% of global market share worlwide, where retail value of luxury goods market worth 339.4 billion dollars and the value of the womenswear industry worth 521 billion dollars [2]. Global Fashion revenue predicts to grow, 33.2% of the fashion enthusiast is 25-34 years old, follows by 23.7% are 34-44 years old and 53.8% of the market dominated by female [3].

According to Jander & Anderson [4], Fashion Week is the most important event of fashion industry aiming to introduce new trends and collections on moving bodies, presenting how the cut and fabric interplayed with the body [5]. A moving body represents by fashion models, a special talent in fashion who displays the attractiveness of the brand in advertisements and runways [1].

In term of searching for fashion models, modeling agency plays a role in looking for new talents, also called scouting. Signing up to a reputable agency can be a gateway to modeling agency for an aspiring model to make a career. Modeling agency is the curators that discover beauty to define aesthetic of agency [6]. Model agencies supply compulsory services for fashion designer by 6.5% of total industry revenue, as models are booked for fashion show, primarily for the show in New York, Paris, Milan and London.The global fashion industry has emerged number of cities worldwide hosting fashion weeks, but the major cities, Paris, New York, Milan, and London, so called "The Big Four" leading the global fashion capital. The "big four" has a long history through the fashion evolution [7].

According to Jeni Rose, Vice President of scouting and development of IMG and Emma Quelch, director of IMG Models on interview with Marie Claire, new talents are frequently found on variety of places such as high street store, festivals, and concerts. for instance, Kate Moss was discovered in JFK airport, Adriana Lima in a charity fashion show. However, the emerging of digital platform, especially Instagram has unfolded the opportunity for potential talents to be discovered through Instagram's hashtag #WLYG.

The growing popularity of social media and online technology has opened a new way for talents to self-promote and agents to scout for new talents. Recently, model agencies outsource talent scouting and other related services that cost 25.3% of revenue [8]. Industry enterprise will continue to deal the internal competition along with the further expands of social media. The number of enterprises increase with average annual rate of 2.5% until 2019. Social media growth exposes a competition among agencies to hire potential top talents as their profile reveals widely online.

Many fields of study have been shifting to digitalization, including fashion industry. One of key pillars in fashion industry 4.0 is the emerging of artificial intelligence [9], allowing us to imitate the knowledge of traditional talent scouting process into an automation model based on machine learning practice. Machine Learning (ML) describes as an automated data analysis process that extract patterns from data through supervised or unsupervised learning [10;11].

Machine Learning and Data Mining is common analytical method used [12]. Machine learning is a branch of artificial intelligence that enables machine to perform the job without programming explicitly by using intelligent software; requiring data to learn in prior [13]. Classification is type of Machine learning algorithm, building a mathematical model containing input and output desired [14]. Several studies use classification techniques such as Decision Tree, and Random Forest to predict the science of success in art and cultural productions [1;15;16].

The primary objective of this study is to build an automation model of talent scouting process using application of Machine Learning for Classification task. Specifically, the paper consider Decision Tree, and Random Forest to predict early success of fashion model under the merging framework of science of success. The accuracy of prediction model depends on the type and structure of data and implementation of feature extraction and engineering give great improvement. The secondary objective is to look at the most suibtable classsification algorithm to predict the success of fashion model and to support decision making in talent scouting.

In accordance with the explanation above, the author is interested to do research entitled Prediction of The Model Prediction for the Next Runway Model using Decision Tree and Random Forest (Case Study: The Big Four Fashion Week)

2. Theoretical Background

2.1 Theoretical Framework

Fashion is used to communicate taste and lifestyle of individual in society where its aesthetic, cultural and social life regularly changes. Fashion show is one of marketing strategies in fashion that mediate business to gain value and enhance image by its unique experiential strategy through fashion show that offers atmosphere building, storytelling, perceptual space and touching experience. [1] Explains that models present the attractiveness of the brand; therefore, fashion model attribute (physical measurement, professional characteristics, and cumulative advantage, in this case social media, instagram are the major consideration. Machine Learning techniques are used to predict the new faces who will be featured for the upcoming fashion show. Techniques used to classify are Decision Tree [17], and Random Forest [18][[19]. The highest accuracy result from the algorithm able to predict the new models presenting in the next fashion week in big four fashion cities before the season start, where the information is valuable for the casting director.

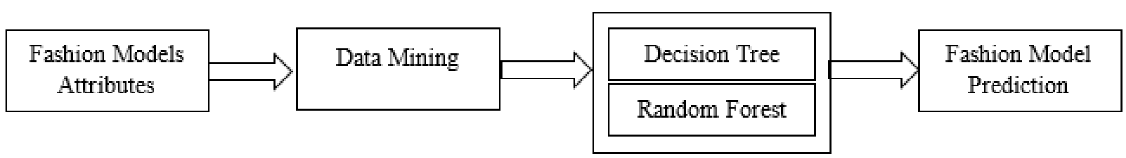

Figure 2.1 Research Framework

2.2 Machine Learning

Machine learning is type of Artificial Intelligence (AI) that provides computer with capability to learn from data without explicitly being programmed [20]. It refers to the changes in system that perform such task involves recognition, diagnosis, planning, robot, control, and prediction, which associated with artificial intelligence (AI). Machine learns the changes of its structure, program, or data (based on its inputs) in such manner that its expected future performance improves [21].

Machine learning is the intersection of Computer Science and Statistics. The statistical learning methods constitute the fundamental of intelligence software that is used to develop the algorithms [17]. The machine requires data to learn in training set to find the relation between input and target variable [20].

2.3 Data Mining

According to Dunham (2002) there are two categories of data mining: predictive data mining and descriptive data mining. Predictive data mining produces the model of the system described by given dataset which results are known. Descriptive data mining identfies pattern or relationship data as a way to explore the data to produce new, nontrivial information based on the available database. The goals of prediction and description are achieved by using data-mining techniques.

2.4 Fashion Modeling Business

According to [22], in fashion modeling business, there are several major components that shape the business: the client, agency, model and consumer. The models play role in representing the product in a way consumers perceive the products. In this cycle, the model seen as the end product to a marketing or advertising idea. They play role in brand communication with the customers, framing consumer experience to lead into consumption [1]. According to [1] studies show that fashion model attributes consider both physical characteristics and subjective consideration.

Physical attractiveness such as weight and body image have understandable importance for social interaction and are important determinants of popularity and influence on others [23]. Professional characteristic is a subjective consideration such as the reputation of the agency representing the model which is an essential criterion for casting [1]

2.5 Talent Management

Talent management is about recruiting, developing and retaining high performance people to grow within the organization [24]. It is an important process to ensure that organization have the quality and quantity of people to meet their current and futures business [25]. However, each organization has different characteristics of talented people that change following the market situation [26].

2.5.1 Talent Selection

Talent selection is one of activities in talent management, defined as a process that choose the eligible candidate among group of applicants [25]. Management focuses on the most potential applicant for the job chosen.

In term of fashion model selection, there are several methods used: interview, network analysis, participant observation, observant participation [27]. Moreover, Instagram has changed the game, where new talent is not only found on the street, but through hashtag and comment. Instagram has evolved into an interactive digital look book for casting new faces [28].

2.6 Decision Tree

Classification is also sometimes called Decision Tree, which its technique assigns categories to a collection of data in order to aide in more accurate predictions and analysis [19]. Decision tree is a data mining methodology applied to solve classification problems. It is a a Hierarchical classifier that applied for supervised learning where the local is identified in a sequence of recursive splits through decision nodes with test function [29]. Decision Tree classifier model predicts the target variable basd on various input variables, in which data can be both continuois and categorical [30]. There are three type of nodes in decision tree: a root node (one incoming edge and zero or more outgoing edges), internal nodes (exactly one incoming edge and two or more outgoing edge), and leaf or terminal nodes (exactly one incoming edge and no outgoing edges). The leaf node is assigned a class label [31].

2.7 Random Forest

Random forest is one of widely used decision tree-based methods in machine learning since its first introduction in Breiman [17]. Random forest is originally an advance development method of CART decision tree using bagging and random feature selection [18]. It is built by combining the prediction of several trees, each of models are trained independently and the predictions are combined through averaging [31].

Random forest is considered as one of the most accurate learning techniques. It is fast and easy to implement, produce highly accurate predictions and can handle a very large number of input variables without overfitting [33]. The algorithm uses the common techniques of boostrap bagging where a random sample from training set is selected to construct trees. As long as tree are not correlated, the procedures lead to better performance while the predictions of a single tree are highly sensitive to noise in its training set [19]. Random forest consists of using randomly selected inputs or combination of inputs at each node to grow each tree. It gives accuracy that has desirable characteristics: accuracy is as good as AdaBoost, relatively robust to outliers and noise, faster than bagging or boosting, giving useful internal estimates of error, strength correlation, and variable importance [18].

2.8 Predictive Analytics

Prediction, also known as forecasting is predicting future event, usually based on historical data by employing statistical model that are computationally fast in forecasting [34]. In fashion, forecasting is finding an important sequence between politics, economics, technology, are, music and therefore fashion to rise a new concept of communication and product [35]. Forecaster will examine the social movements to predict because society creates impacts in trend. People has logical choices based on established principles and concepts of fashion [36]. According to Dunham (2002), in data mining, prediction produces the model of the system described by given dataset which results are known.

3. Methodology

Table 3.1 Research Characteristics

No	Research Characteristic	Type
1.	Based on Method	*Quantitative*
2.	Based on Purpose	Descriptive
3.	Based on Researcher Involvement	No Data Intervention
4.	Based on Analysis Unit	Individual/ Social Media
5.	Based on Time	*Cross sectional*

Source: Author Documentation

3.1 Data Collection

In this research, there are 1034 profiles collected: 312 new faces, and 522 appear on 2018 Spring/Summer Fashion Week. Moreover, there are 13 variables: 10 continuous data and 3 binary data. There are three main characteristics inclueded in this study: physical characterics, professional charachteristics, and social influence characteristics. Each characteristics consist of following attributres:

- Physical characteristics: height, bust, waist, hips, dress, and shoes.
- Professional characteristics: agency score, magazine covers, editorial, and advertisement.
- Social influence characteristics: number of posting, number of likes, and number of comment.

3.2 Data Pre-processing

the data cleaning works to clean the data by filling in missing values, smoothing noisy data, and transforming to standardize the data. The raw data is transformed into more structured data. In this step, unnecessary data is eliminated, missing value is replaced by median, each data is labelled, and duplicate data is removed. The aim is to minimize error at the processing step.

3.3 Data Processing

After the preprocessing, the data is split into 70:30, where 70% data for training and 30& data for testing. Training data is to practice the model using labelled data and testing data to validate the accuracy of the model.

Data is processed by building three models of data. Each model consists of different combination of defining variables. The model constructions are mentioned as below:

- Model I: Physical Characteristics
- Model II: Physical + Professional Characteristics
- Model III: Physical + Professional + Social Media Characteristics

The implementation uses Rapid Miner machine learning tool by optimally tuning the parameters as follows:

- Measure the quality of decision splits by using entropy,
- Maxax-depth to 10 for Decision Tree and Random Forest,
- Adopting a maximum of 100 trees for Random Forest,

- Evaluation uses F-measure and accuracy score,
- K-Fold Cross Validation (k=10).

The outcome expected is a prediction of becoming popular (appear in 2019 A/W Fashion Week) or not becoming popular in the future fashion event

3.4 Evaluation and Validation

In order to minimize biases associated with random sampling of training data samples to compare the prediction accuracy of two or more methods, which may use a methodology called k-fold cross validation. In this research, the results obtained by averaging one thousand iterations of *k*-Fold Cross Validation (*k* = 10).

The measurement of how accurate the classifier is at the predicting the class label by evaluating its accuracy (recognition rate), sensitivity (recall), and precision. Accuracy also refers to a classifier's predictive abilities (Han, Kamber, & Pei, 2012: 364). Primary source of performance measurement for classification problem is coincidence matrix, also known as classification matrix or a contigency table.

4. Result and Analysis

Data that has been through data preprocessing remaining 764 data. It is collected from three main sources: Fashion Model Directory for physical characteristics and professional characteristic, Models.com for rating of modeling agency, and Instagram for social media characteristics. The frequency of model's appearance on Fashion Week is as below:

- Label YES : 460 profile data
- Label NO : 303 profile data

4.1 Decision Tree Result

Decision Tree classification technique is conducted to predict the talents to be featured on the next season of Fashion Week. Each model construction represents different result as shown below.

Table 4.1 Decision Tree Confusion Matrix

	True Positive	***False Positive***	***True Negative***	***False Negative***
Model I	173	75	58	44
Model II	179	75	58	38
Mode III	180	66	67	37

Source: Author Documentation

According to the above, Decision Tree performs best with combination set of data on Model III. Decision Tree is able to predict 180 out of 246 fashion model who appear on A/W Fashion Week 2019 correctly, and 67 out of 104 not popular fashion model precisely. Total valid prediction are 247 out of 350 data, leaving 29.5% of data misclassified.

Table 4.2 Decision Tree Performance Measurement

	Accuracy	***Precision***	***Recall***	***Classification Error***	***F-Measure***
MODEL I	66	70.71	79.85	34	74.35
MODEL II	67.71	71.55	82.68	32.29	75.49
MODEL III	70.57	73.91	83.05	29.43	77.65

Source: Author Documentation

Model III scores the highest accuracy of 70.57% and the highest precision of 73/91% and recall of 83.05%. Moreover, it performs 77.65% F-measure with lowest classification error 29.43%%.

4.2 Random Forest Result

Random Forest classification technique is conducted to predict the talents to be featured on the next season of Fashion Week. Each model construction represents different result as shown below.

Table 4.2 Random Forest Confusion Matrix

	True Positive	***False Positive***	***True Negative***	***False Negative***
MODEL I	190	89	44	27
MODEL II	201	76	57	16
MODEL III	205	77	56	12

Source: Author Documentation

According to the above, Random Forest performs best with combination set of data on Model III. Decision Tree is able to predict 205 out of 282 fashion model who appear on A/W Fashion Week 2019 correctly, and 56 out of 68 not popular fashion model precisely. Total valid prediction are 261 out of 350 data, leaving 25.4% of data misclassified

Table 4.4 Random Forest Performance Measurement

	Accuracy	***Precision***	***Recall***	***Classification Error***	***F-Measure***
MODEL I	66.86	68.46	87.58	33.14	76.64
MODEL II	73.71	73.35	92.66	26.29	81.46
MODEL III	74.57	73.20	94.52	25.43	82.24

Source: Author Documentation

The result of the performance of Random Forest is shown on table 4.4. Model III scores the highest accuracy of 74.57% and the highest precision of 73.20% and recall of 94.52%. Moreover, it performs a good F-measure 82.24 with lowest classification error 25.43%.

4.3 Model Comparison

This study conducts three different models. Additional variables put to each model before employs Decision Tree and Random Forest for prediction. Additional variables give positive impact to the accuracy and F-measure; the incremental indicates better prediction model. Decision Tree raises accuracy with average 2.28% of each model. Random Forest also points an increment significantly from each model by 3.85% as shown below.

Table 3.5 Prediction Model Comparison

	MODEL I		**MODEL II**		**MODEL III**	
	Accuracy	F-measure	Accuracy	F-measure	Accuracy	F-measure
Random Forest	66.86	76.64	73.71	81.46	74.57	82.24
Decision Tree	66	74.35	67.71	75.49	70.57	77.65

Source: Author Documentation

The table above indicates that additional model improves the machine performance to predict the next fashion model for the upcoming show. Social media activities are a prominent key within fashion industry, along with their professional records and the aesthetics to represent the art of designers' work.

This study constructs three different models on Decision Tree and Random Forest. These learning techniques are tree-based that behaves as "if this than that" to decide. As explained above, the construction of Model III consists of physical, professional and social media characteristics score the best. In order to see which classification techniques that performs better, we compare the two learning techniques.

Random forest impressively scores accuracy 74.57% with lowest classification error 25.43% and F-measure 82.24%. The performance of Random Forest correctly predicts 261 out of 350 of total sample on testing data.

Table 4.6 Performance Measurement Comparison

	Accuracy	***Precision***	***Recall***	***Classification Error***	***F-measure***
Decision Tree	70.57	73.91	83.05	29.43	77.65
Random Forest	24.57	73.20	94.52	25.43	82.24

Source: Author Documentation

Decision Tree works impressively fast and easy to interpret; however, Random Forest builds and merges multiple decision tree in order to get an optimal choice at each node and a more stable and accurate prediction.

4.4 Descriptive Analysis

Physical characteristics is the first variable analyzed in this study. The distribution of models' physical characteristics is available on table 4.7. Comparing the models body measurement to US female population for group of age 20 and over, fashion model is much taller; as matter of fact, the shorted fashion model is still taller than average US female population, 161 cm [32]. In general, their body measurement is the contrary from population, such as average models' waist is smaller than US female population, 98cm.

Table 4.7 Summary of Fashion Models Physical Characteristics

	Height	***Bust***	***Waist***	***Hips***	***Dress***	***Shoes***
Mean	178.06	79.33	60.04	87.47	34.38	39.44
Mode	178	81	61	89	35	39
Minimum	165	31	24	34	30	36
Median	178	80	60	88	35	39
Maximum	186	91.5	88	110	40	42
Standard Deviation	2.12	3.41	2.83	2.74	0.94	1.12

Source: Author Documentation

Adding the information on professional characteristics, the author finds a strong association that being hired on top agency helps to rise changes to be featured for future works. Moreover, social media characteristics, particularly Instagram likes, helps to increase the chance of fashion models' success.

5. Conclusion and Suggestion

5.1 Conclusion

The trend of social media sharing platform, in this case Instagram, has increased the predictability accuracy by 0.86% in Random Forest and 3.4% in Decision Tree. Therefore, social media can give a signal in selecting fashion models. In this research, Random Forest can impressively predict 261 out of 350 correct prediction with 74.57% highest accuracy and 82.24% F1 score. Furthermore, Random Forest performances 25.43% classification error.

Moreover, this research found that in Random Forest, the importance factors in selecting runway model are Instagram likes, and appearance on advertisement, editorial and magazine cover. Social media and professional characteristics found important to be considered.

5.2 Suggestion

Fashion brand reputation is an important aspect. Future research is suggested to treat the brands differently; models who run for high prestige brand might have higher weight than new or relatively unpopular brands. Including more social media platform such as Facebook, Twitter, and YouTube expected to reveal more insight about the impact of social media in talent recruitment and selection process, especially in visual centric business. The emerging social media platform, in this case Instagram gives greater insight about the importance of social influence and popularity of fashion models to the chance of getting selected by modeling agency. Besides, utilizing social media information might help modeling agency to reduce scouting cost.

This study is limited to analyze female fashion, however male modeling is increasingly becoming more popular. Fashion practitioner should actively create a database for male models as it is still limited. Furthermore, the same study can be applied for male models to have interesting insight on male contribution to the female dominated industry.

References

[1] Park, J., Ciampaglia, G. L., & Ferrara, E. (2016). Style in the Age of Instagram: Predicting Success within the Fashion Industry using Social Media. *CSCW*, 64-73.

[2] Adegeest, D.-A. (2018, January 12). *Global fashion sales to increase in 2018, says McKinsey report*. Retrieved from Fashion United: https://fashionunited.uk/news/business/global-fashion-sales-to-increase-in-20218-says-mckinsey-report/2018011227622

[3] Statista. (2018, December 24). *Fashion*. Retrieved from Statista: https://www.statista.com/outlook/244/100/fashion/worldwide#market-globalRevenue

[4] Jandér, L., & Andersson, V. (2016). Social Media, Insta-Culture & Reinvention of Fashion Week. *Thesis for One-Year Master, 15 ECTS Textile Management*.

[5] Skov, L., Skjold, E., Moeran, B., Larsen, F., & Csaba, F. F. (2009). Creativity at Work: The Fashion Show as an Art Form. *Creative Encounters Working Paper*.

[6] Marie Claire. (2015, April 21). *What Model Scouts are Looking for (and The New Place to Get Spotted)*. Dipetik June 25, 2019, dari Marie Claire: https://www.marieclaire.co.uk/news/fashion-news/what-model-scouts-look-for-89318

[7] Global Fashion Capital. (2015). *Global Fashion Capital: A Teaching Guide.* The Museum at FIT.

[8] Robehmed, N. (2017). *Highest-Paid Models 2017: Kendall Jenner Takes Crown From Gisele with $22M Year*. Dipetik June 25, 2019, dari Forbes: https://www.forbes.com/sites/natalierobehmed/2017/11/21/highest-paid-models-2017-kendall-jenner-ousts-gisele-with-22-million/#7249a7e12b5c

[9] Lectra. (2018). *The Rise of Industry 4.0 in Fashion.* Lectra.

[10] Kelleher, J. D., Namee, B. M., & D'Arcy, A. (2015). *Fundamentals of Machine Learning for Predictive Data Analytics: Algorithms, Worked Examples, and Case Studies.* Massachusetts: The MIT Press.

[11] Arusada, M. D., Putri, N. A., & Alamsyah, A. (2017). Training Data Optimization Strategy for Multiclass Text Classification. *International Conference on Information and Communication Technology*, 1-5.

[12] Alpaydin, E. (2010). *Introduction to Machine Learning.* Cambridge: The MIT Press.

[13] Mohammed, M., Khan, M. B., & Bashier, E. M. (2017). *Machine Learning Algorithms and Application.* Boca Raton: CRC Press.

[14] Russel, S. J., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach.* New York: Pearson Education.

[15] Arakelyan, S., Morstatter, F., Martin, M., Ferrara, E., & Galstyan, A. (2018). Mining and Forecasting Career Trajectories of Music Artists. *ACM Conference on Hypertext and Social Media*, 9.

[16] Ren, J., Shen, J., & Kauffman, R. J. (2016). What Makes A Music Track Popular in Online Social Networks? *International Conference Companion on World Wide Web*, 95-96.

[17] Au, T. C. (2018). Random Forest, Decision Tree, and Categorical Predictors: The "Absent Level" Problem. *Journal of Machine Learning Research 19*, 1-30.

[18] Breiman, L. (2001). Machine Learning Random Forest.

[19] Ponmani, S., Samuel, R., & Vidhupriya, P. (2017). Classification Algorithms in Data Mining - A Survey. *International Journal of Advance Research in Computer Engineering & Technology*, 1.

[20] Budiharto, W. (2016). *Machine Learning & Computational Intelligence.* Yogyakarta: ANDI.

[21] Nilson, N. J. (1998). *Machine Learning: .* Stanford: Standford University Press

[22] Shamley, T. (2011) *The Business of Fashion Modeling.* US: Dhymon House Publishing

[23] Anyzova, P., & Mateju, P. (2018). Beauty Still Matter: The Role of Attractiveness in Labor Market Outcomes. *International Sociology*, 269-291.

[24] Meyers, M. C., Van Woerkom, M., & Dries, N. (2013). Talent- innate or acquired? Theoretical Consideration and their Implications for Talent Management. *Human Resource Management Review 23 (4)*, 305-321.

[25] Abdollahbeigi, B., Salehi, F., & Jayashree, S. (2017). The Effect of Recruitment, Selection and Devemopment on Talent Managemen in IKCO Company in Iran. *International Journal of Advance Engineering and Management*, 69-77.

[26] Samuel, L. (2009). From Talent Compliance to Talent Commitment. *Stategic HR Review vol. 6*, 32-35.

[27] Mears, A. (2011). *Pricing Beauty: The Making of a Fashion Model.* Oakland: University of California Press.

[28] Style, C. i. (2016, May 23). *highscnobiety*. Retrieved from highscnobiety: https://www.highsnobiety.com/2016/05/23/instagram-modeling/

[29] Kantardzic, M. (2011). *Data Mining Concepts, Models, Methods, and Algorithm.* Canada: John Wiley & Sons.

[30] Guggari, S., Kadappa, V., & Umadevi, V. (2018). Non-sequential Partitioning Approach to Decision Tree Classifier. *Future Computing and Information Journal 3*, 275-285.

[31] Tan, P. N., Steinbach, M., & Kumar, V. (2005). *Introduction to Data Mining.* Massachussets : Pearson Education Inc.

[32] Denil, M., Matheson, D., & Freitas, N. d. (2014). Narrowing the Gap: Random Forest in Theory and in Practice. *Proceedings of the 31st Machine Learning Resesarch* (pp. 665-673). PMLR.

[33] Biau, G. (2012). Analysis of a Random Forest Model. *Journal of Machine Learning Research*, 1063-1095.

[34] Yu, Y., Choi, T.-M., & Hui, C.-L. (2011). An intelligent fast sales forecasting model for fashion products. *Expert System with Application*, 7373-7379.

[35]Fimmano, P. (2019, january friday). *Polimoda*. Retrieved from polimoda: https://www.polimoda.com/courses/fashion-trend-forecasting

[36] Rilley, A. (2014). *Key Concepts for the Fashion Industry.* London: Bloomsbury.