# Sales Potential: Modelling Sellability of Visual Aesthetics of a Fashion Product

Vikram Garg
Myntra
Bengaluru, India
vikram.garg@myntra.com

Rajdeep H Banerjee
Myntra
Bengaluru, India
rajdeep.banerjee@myntra.com

Anoop K R
Myntra
Bengaluru, India
anoop.kr@myntra.com

Sreenivas T
Myntra
Bengaluru, India
sreenivas.t@myntra.com

Deepak Warrier
Myntra
Bengaluru, India
deepak.warrier@myntra.com

## ABSTRACT

Fashion is a domain primarily driven by its visual content. Look and feel of a fashion product is difficult to quantify as it is essentially subjective and driven by a host of subtle factors. In this work we formulate a mechanism that grades the look of a product, which we call Sales Potential (SP), that captures visual aesthetics. Our approach normalizes the effects of merchandising factors like discounts, price, list views and brand effects introduced in the e-commerce platform that influences buyers behaviour. Our approach also relies on the observation that similar looking fashion products(styles) should have similar scores in a given fashion category. We show two applications of this SP score in assortment planning. In one we look for products in other social/e-commerce platforms and grade them relative to our platform. In second we look at grading products within the platform for replenishment.

## KEYWORDS

Sales Potential, Similarity, Merchandising, Top Seller, Design

**ACM Reference format:**
Vikram Garg, Rajdeep H Banerjee, Anoop K R, Sreenivas T, and Deepak Warrier. 2016. Sales Potential: Modelling Sellability of Visual Aesthetics of a Fashion Product . In *Proceedings of , , ,* 6 pages.
DOI: 10.1145/nnnnnnn.nnnnnnn

## 1 INTRODUCTION

Fashion is a domain strongly associated with its visual content [5, 6]. Characterizing a fashion product is complex due to large variability in customer perception arising due to demography, trends etc. Discerning visual aesthetics which sell well is critical in a fashion e-commerce space like Myntra to address buying, substitution and replenishment. [1]

[1]Myntra is the largest online fashion retailer in India with 18M+ monthly average users and around $0.3M$ products live


DOI: 10.1145/nnnnnnn.nnnnnnn

Differentiating a fashion product as being good or bad based on visual content alone is an ill-posed task. Manual grading of a fashion product is very subjective and not scalable. A fashion product/style, henceforth used interchangeably, has many soft attributes and requires to be understood holistically in terms of style, fabric, visual features, brand etc. Traditionally in fashion a product is graded based on its platform merchandising values (e.g. Gross Margin, Revenue, Quantity sold) alone and doesn't consider its visual aspects. In this work we address this gap and attempt to formalize "Sellability" of a fashion product in terms of its visual aspects and refer to this as "Sales Potential" (SP). This formulation normalizes the bias introduced in platform because of merchandising effects such as MRP, discounting, visibility etc and relates only to visual aesthetics (trendiness, design, feel etc.). We further investigate the following two approaches to understand the effect of product images. First approach constrains the above formulation with visual similarity of style. And the second approach uses a regression model on deep learned features of product images to obtain its sellability. We perform experiments to analyse both these approaches against a standard Depricing model [1].

As an application of SP we look at product assortment which is a key differentiator across e-tailers and serves to improve overall customer experience. Grading the assortment helps in timely replenishment/substitutions and buying decisions. Also discovering fashion products on other social/e-commerce platforms that have high SP on our platform enables better assortment planning and improves freshness. We provide a framework wherein the regression model is used to compute SP and subsequently grade images from other fashion platforms.

Contributions of this work are as follows:

- Ideation and development of a novel formulation of Sales Potential (SP) in fashion using visual content.
- A framework to grade styles within and across other e-commerce/social platforms.

## 2 RELATED WORK

Quantifying visual aesthetics is a difficult problem. Impact of images on sales in ecommerce is well studied in [5]. However, most work in literature uses visual information only for attributing users preferences to an item. In [10, 11] a holistic style recommendation model over parts of the image is learnt. Given a query image like top wear they suitably return relevant images of bottom wear that is closest to the query based on the learnt model. In [7, 9] a personalized recommendation system for users based on visual features and item affinity is addressed. A time aware user recommender system for items affinity are built in [15]. A mixture of time aware and visual aware recommender modelling is done in [8].

As mentioned all the above works try to address some form of user recommendation without explicitly addressing selling potential of a fashion product alone. In [8] a Depricing model is suggested that smooths the effect of price biases which is used in retail demand forecasting. However, this doesn't address any visual aspects of the style. A description of fashion items based on non visual attributes like color, fabric, price buckets are described in [3]. Purchase probabilities of the product is estimated based on customer to product recommendations. But they do not normalize for the bias introduced by the platform in terms of discount, price, visibility and brand affinities etc.

In this work we attempt to address the gap between visual aspects and sellability of the fashion product on a fashion e-commerce platform by normalizing the various merchandising factors like MRP, discounts, visibility and brand effects introduced by the platform.

## 3 SALES POTENTIAL

Sellability of a fashion product is a characteristic of attributes of the product, how it looks/feels and what a customer thinks about it. However, on an e-commerce platform this is often biased by intrinsic factors like brand, MRP, discounts , listcounts etc. It is also influenced by extrinsic factors like user preferences, seasonality etc. which are non product related. "Sales potential (SP)" is a score that captures the visual feel and proportionally relates customer preferences of a fashion product in a holistic sense after normalizing for merchandising factors (MRP, discounts , listcounts) and brand. This is formalized as:

$$\text{SP} \propto \textsc{Indicators} \tag{1}$$

Where Indicators could be Quantity(Q) or Revenue. In this work we restrict the scope of indicators to quantity sold. Also, the above definition limits itself to a time period, wherein the indicators may vary from season to season.

### 3.1 Indicators/Features

To set the context for formalizing SP we define what are sales indicators and the features which influence them.

*Indicators.*

- Quantity sold: We look at the total sales of a product in a category (e.g.: Mens T-Shirts) in a 3 month period, considering only those styles which are live on the platform for a minimum of 30 days.
- Revenue: The total currency earned by sales of the product in the considered time duration.
- CTR: Is the average click through rate, defined as the ratio of the number of clicks on the product page and total listings of the product.

*Features.*

- Average selling price (ASP): Price after accounting for discount of the product averaged over a time window. This is an important correction as Myntra operates in a predominantly discount driven market.
- Listcounts (Visibility): The visibility of a product is the listings a product gets. Higher up the product in the list page on the platform the more list views it gets. This is a bias in the system which promotes some products over others making them more easily discoverable to the shopper.
- Brand: The product brand creates bias because some shoppers are inherently biased more towards certain brands. This leads to very different sales numbers for 2 very similar looking products but belonging to different brands.
- MRP: Maximum Retail Price. This refers to actual price of the product. This plays a major role in customer buying preferences.

Customers engage with different products differently on the platform, and how they engage define the sellability of that product. Quantity sold and revenue for a product can be interpreted as directly capturing SP of style. Other indicators like click-through-rate (CTR) are more nuanced indicators of SP. For a particular category e.g. men's shirts, a rank ordering of any one of the sales indicators would give a fair indication of the sellability of the products in that category.

However customer engagement with products is heavily influenced by the biases built in the system, in the form of discounts, product visibility and product brand. In such a case interpreting a higher value for indicators as more sellable is a distorted measure; to get a clearer measure we attempt to normalise for the biases in the system. Formally, we try to derive a single number score which tells where a given product stands relative to other products in the same category with respect to sellability in the next section.

## 4 SP FORMULATION

Let $f_i$ be the features. Here features are merchandising factors affecting the style. These are the factors which are set for a style based on past knowledge. These factors are in the control of the e-commerce platform. Let the quantity sold of the fashion product be $q_i$. To account for the biases (merchandising + brand) as stated above we introduce normalizing parameter $\alpha_i$ on each of these features. Hence we extend our formulation in 1 for a style $j$ using power law ( accounts for large scale variations) as in depricing model [8] below

$$\left(\prod_{i=n}^{d} f_{ij}^{\alpha_i}\right) SP_j = Q_j \tag{2}$$

Taking logarithms on both sides of above equation 2 we obtain

$$\left(\sum_{i=n}^{d} \alpha_i f_{ij} + log(SP_j)\right) = log(Q_j) \tag{3}$$

At this point we try to bring in visual aspects using two approaches. Since styles are normalized on merchandising factors it is natural to expect similar looking styles to have similar SP scores. Hence in the first approach we impose visual similarity constraints to obtain the below objective

$$\min \gamma_{jk} \tag{4}$$

where $j$ and $k$ are visually similar looking styles and $\gamma_{jk} = |log(SP_j) - log(SP_k)|$. Henceforth we shall refer this approach as **Proposed**

**SP (P-SP)**. Substituting equation 3 in objective 4 we get

$$\min \sum_{j,k} \left| \left( |\log(Q_j) - log(Q_k)| + \sum_{i=1}^{d} \alpha_i |f_{ij} - f_{ik}| \right) \right| \quad \forall j,k \; similar \tag{5}$$

where $d$ is the number of features. The above Least Absolute Deviations (LAD) [2] objective is then solved for $\alpha_i$. Solved parameters are substituted in equation 3 to obtain P-SP.

In the second approach we directly associate the SP scores to images features $f_{vgg}$ obtained from fully connected layer of a VGG-16 network [14] pretrained using Imagenet [4] as below

$$\log \text{DL-SP} = \beta^{\intercal} f_{vgg} \tag{6}$$

where $\beta$ is the weighting parameters to be learnt. This is substituted in equation 3 and is solved using regression. We call this **Deep Learnt SP (DL-SP)**.

Intuitively SP latently represents the look and feel of the fashion product. It also captures customer demands on the platform after normalizing for all its biases.

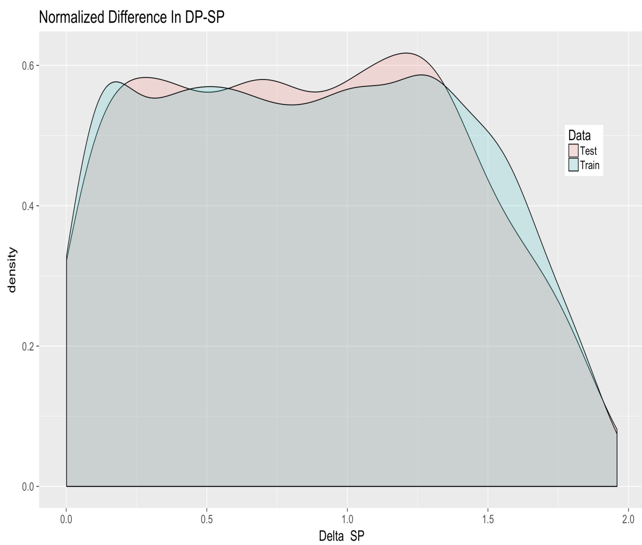


**Figure 1: ΔDP-SP differences on train and test data.**

## 5 SP ANALYSIS

We analyse P-SP, DL-SP with an existing industry model for depricing [8] which we call as **DP-SP** in equation 7. DP-SP forms our baseline.

$$\text{DP-SP} = \frac{revenue}{log(1 + Listcounts/1000)} \left( \frac{ASP}{MRP} \right)^2 \tag{7}$$

We compare the percentage difference of all the three SP scores based on their average for similar styles using equation 8.

$$\Delta SP_{ij} = \frac{|SP_i - SP_j|}{(SP_i + SP_j)/2} * 100 \tag{8}$$

In this work we have used sales data on Myntra platform for Men T-shirts category. The assortment (~60000 styles) consist of 100+ brands listed on Myntra and which were sold over a three month period. We used similarity identification [13] algorithm to identify ~27k similar styles pairs. We split these pairs in 3:1 ratio to obtain training and testing sets for our experiments.

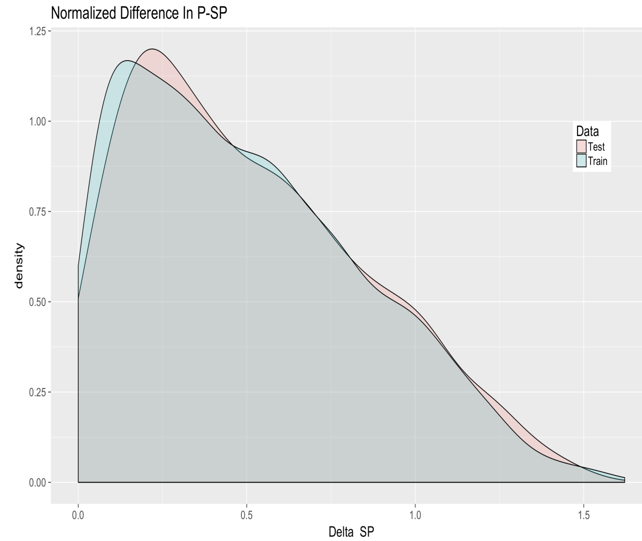


**Figure 2: ΔP-SP differences on train and test data.**

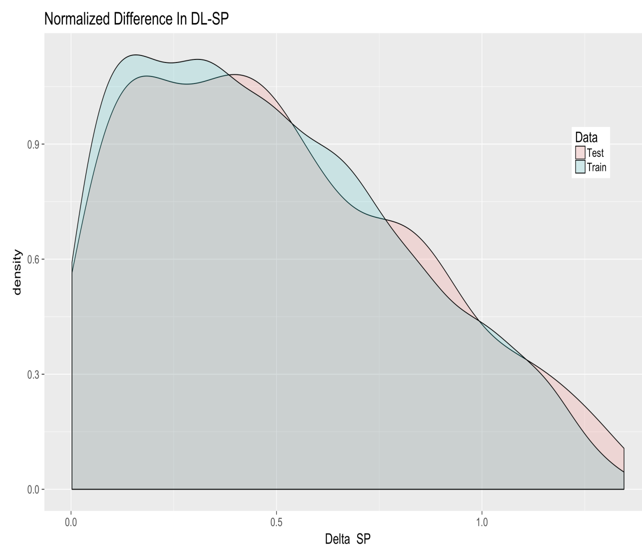


**Figure 3: ΔDL-SP differences on train and test data.**

Figures[1,2,3] shows the plot of $\Delta SP$ between all similar pairs of styles. From the plots we observe that ΔDP-SP has a near uniform distribution. The ΔDL-SP and ΔP-SP exhibit better peaky behaviour. This indicates the DP-SP is agnostic to visual aspects of styles. However, ΔDL-SP has more variance than ΔP-SP. This may be because DL-SP falls short of capturing visual features, through a pretrained model, which holistically define a style . This probably can be circumvented by building an end-to-end system which normalizes merchandising factors and also learns visual features which can comprehend style with some design stimuli.

| ΔDL-SP : 36.54<br>ΔDP-SP : 81.60<br>ΔP-SP : 0.84 |  |  |
|---|---|---|
| style_id | 1491665 | 1451404 |
| **DL-SP** | **0.7** | **0.484** |
| **De-Priced-SP** | **4055.292** | **1705.044** |
| **ProposedSP** | **0.05** | **0.05** |
| Quantity | 11 | 9 |
| MRP | 1899 | 799 |
| ASP | 1819.24 | 749.88 |
| ListCounts | 91655 | 31671 |
| Discount | 0.042 | 0.061 |

**Figure 4: Similar looking style pair with same Base-color but differing in ASP, Brand and Visibility.**

| ΔDL-SP : 149.03<br>ΔDP-SP : 8.75<br>ΔP-SP : 8.72 |  |  |
|---|---|---|
| style_id | 1460568 | 1460566 |
| **DL-SP** | **0.326** | **2.234** |
| **De-Priced-SP** | **20963.801** | **19207.17** |
| **Proposed-SP** | **0.136** | **0.149** |
| Quantity | 95 | 102 |
| MRP | 1599 | 1199 |
| ASP | 1492.08 | 1149.93 |
| ListCounts | 359518 | 274091 |
| Discount | 0.067 | 0.041 |

**Figure 5: Similar looking style pair with same Brand and Base-color but differing in ASP and collar.**

We also validated SPs with the CTR values of the styles in our platform as show in Table 1. Pearson Correlation provides a clear validation that the proposed SP is better at imitating the customer behaviour in terms of CTR. We also measure the Maximum Information Coefficient [12] in the same Table 1. Proposed SP provides the best MIC which further boosts the importance of the proposed approach.

| ΔDL-SP : 64.14<br>ΔDP-SP : 58.27<br>ΔP-SP : 31.71 |  |  |
|---|---|---|
| style_id | 1509812 | 1161647 |
| **DL-SP** | **1.181** | **0.607** |
| **De-Priced-SP** | **9317.947** | **16980.531** |
| **ProposedSP** | **0.188** | **0.136** |
| Quantity | 145 | 135 |
| MRP | 899 | 899 |
| ASP | 647.64 | 828.4 |
| ListCounts | 185861 | 267323 |
| Discount | 0.28 | 0.079 |

**Figure 6: Similar looking style pair with same Brand and Base-color but differing in ASP.**

| ΔDL-SP : 132.89<br>ΔDP-SP : 133.89<br>ΔP-SP : 0.18 |  |  |
|---|---|---|
| style_id | 1473514 | 1459407 |
| **DL-SP** | **0.141** | **0.701** |
| **De-Priced-SP** | **5813.887** | **1151.157** |
| **ProposedSP** | **0.081** | **0.082** |
| Quantity | 12 | 29 |
| MRP | 3599 | 699 |
| ASP | 3055.76 | 425.81 |
| ListCounts | 93333 | 52553 |
| Discount | 0.151 | 0.391 |

**Figure 7: Similar looking style pair with same Base-color but differing in ASP, Brand and Visibility**

**Table 1: SP with CTR correlations**

| SP | Pearson Correlation | MIC |
|---|---|---|
| DP-SP | −0.04545 | 0.08597 |
| DL-SP | 0.16870 | 0.10121 |
| P-SP | 0.29033 | 0.14294 |

In Figure[4,5,6,7] we show four pairs of visually similar styles. These pairs have varying merchandising values as observed on our platform. In Figure 4 we observe that owing to different revenues these similar looking styles gets very different DP-SP score with normalized ΔDP-SP being approximately 82%. Using Deep learned image based features of the two styles, we achieve a slightly better ΔDL-SP 37%. But using the power law model and applying the image similarity constraint, the P-SP thus obtained, perform the best with ΔP-SP 0.84%. Similarly in Figures [5,6,7] P-SP has significantly lesser percentage difference compared to DP-SP and DL-SP. This indicates that imposing similarity constraint better normalizes for the biases inherent in the platform.

# 6 APPLICATIONS

## 6.1 Cross Platform Image Grading

With a large number of customers on platform fashion product assortment also needs to be scaled and well curated. In order to obtain rich diversity and freshness of products it is necessary to refresh the catalogue with fashion products that may be trending in different social/e-commerce platforms. Thus grading these products on our platforms can help in substitutions and buying decisions during any sale event. However on these platforms we can only find image data without any merchandising factors. Hence we use DL-SP to grade these images within our platform. Our approach is outlined in Figure 8. Here we first extract features of the final fully connected layer by passing the images through a VGG-16 network. We then use our DL-SP model shown in equation 6 to obtain the SP scores.

We took a set of ∼1000 images from Flipkart(Platform1) and ran our DL-SP model on them. Based on our platform threshold we identified top 560 styles from this list. Out of these, 298 similar styles (about 53%) in Myntra turned out to be top sellers on our platform. We show 3 such styles along with similar styles from Myntra(Platform2) in Figure 9. The first two pairs have high SP scores and are for assortment planning on our platform(Myntra). Product in Figure 9(a) on platform 2 has a high quantity of images sold and less discounting. Whereas product in Figure 9(c) on platform 2 had lesser quantity of images sold which is also reflected in their DL-SP scores.

## 6.2 Catalogue replenishment

We also address another important application of identifying 'top sellers' for replenishment within the catalogue. Here we group styles based on certain attributes in a category(like pattern, collar in Tshirts) and within each group find products which are top sellers and bottom sellers based on the P-SP scores. The various product grouping obtained as Top sellers are shown in Figure 10

# 7 CONCLUSIONS

In this work we have attempted to address the Sales Potential of a fashion product in terms of its visual aesthetics. We have provided a formulation for estimating SP of a product by normalizing for the merchandising values and brand which introduces an inherent bias in its sales on the system. The above normalizing coefficients affecting SP of a style is then solved using Image similarity constraints (P-SP). We have shown that P-SP and DL-SP provide better SP score consistency on similar image pairs than a DP-SP retail model which is agnostic to visual aspects. We then show that P-SP score provides a significant correlation with CTR on the platform which captures some visual aspects of the product. We also model the SP by regressing with image features (DL-SP). We show that the P-SP and DL-SP approaches significantly outperform standard depricing model DP-SP (baseline) in terms of CTR correlations on image pairs. The analysis also shows that VGG-16 visual features may not capture the complete context of a style . Further, imposing similarity constraints in DL-SP formulation makes the model complex as the difference of image features may not be a good measure of similarity. Also the visual features needs a better articulation of the product which need to be learnt along with other context like trends, demography, customer segments etc. which we leave for future work. We finally show how the SP score is used in grading products in other e-commerce/social platform and for replenishment within the platform enabling better assortment planning.

# REFERENCES


[1] 2014. Oracle Retail Demand Forecasting. *http://docs.oracle.com/cd/E75759_01/rdf/pdf/1401/rdf-1401-02-fcug.pdf* (2014).

[2] Peter Bloomfield and William Steiger. 1980. Least absolute deviations curve-fitting. *SIAM Journal on scientific and statistical computing* 1, 2 (1980), 290–301.

[3] Christian Bracher, Sebastian Heinz, and Roland Vollgraf. 2016. Fashion DNA: Merging Content and Sales Data for Recommendation and Article Mapping. *arXiv preprint arXiv:1609.02489* (2016).

[4] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on.* IEEE, 248–255.

[5] Wei Di, Neel Sundaresan, Robinson Piramuthu, and Anurag Bhardwaj. 2014. Is a picture really worth a thousand words?:-on the role of images in e-commerce. In *Proceedings of the 7th ACM international conference on Web search and data mining.* ACM, 633–642.

[6] Anjan Goswami, Naren Chittar, and Chung H Sung. 2011. A study on the impact of product images on user clicks for online shopping. In *Proceedings of the 20th international conference companion on World wide web.* ACM, 45–46.

[7] Ruining He, Chunbin Lin, Jianguo Wang, and Julian McAuley. 2016. Sherlock: sparse hierarchical embeddings for visually-aware one-class collaborative filtering. *arXiv preprint arXiv:1604.05813* (2016).

[8] Ruining He and Julian McAuley. 2016. Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering. In *Proceedings of the 25th International Conference on World Wide Web.* International World Wide Web Conferences Steering Committee, 507–517.

[9] Ruining He and Julian McAuley. 2016. VBPR: visual bayesian personalized ranking from implicit feedback. In *Thirtieth AAAI Conference on Artificial Intelligence.*

[10] Tomoharu Iwata, Shinji Wanatabe, and Hiroshi Sawada. 2011. Fashion coordinates recommender system using photographs from fashion magazines. In *IJCAI*, Vol. 22. Citeseer, 2262.

[11] Vignesh Jagadeesh, Robinson Piramuthu, Anurag Bhardwaj, Wei Di, and Neel Sundaresan. 2014. Large scale visual recommendations from street fashion images. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM, 1925–1934.

[12] David N Reshef, Yakir A Reshef, Hilary K Finucane, Sharon R Grossman, Gilean McVean, Peter J Turnbaugh, Eric S Lander, Michael Mitzenmacher, and Pardis C Sabeti. 2011. Detecting novel associations in large data sets. *science* 334, 6062 (2011), 1518–1524.

[13] Devashish Shankar, Sujay Narumanchi, HA Ananya, Pramod Kompalli, and Krishnendu Chaudhury. 2017. Deep Learning based Large Scale Visual Recommendation and Search for E-Commerce. *arXiv preprint arXiv:1703.02344* (2017).

[14] Karen Simonyan and Andrew Zisserman. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).

[15] Yongfeng Zhang, Min Zhang, Yi Zhang, Guokun Lai, Yiqun Liu, Honghui Zhang, and Shaoping Ma. 2015. Daily-aware personalized recommendation based on feature-level time series analysis. In *Proceedings of the 24th international conference on world wide web.* ACM, 1373–1383.

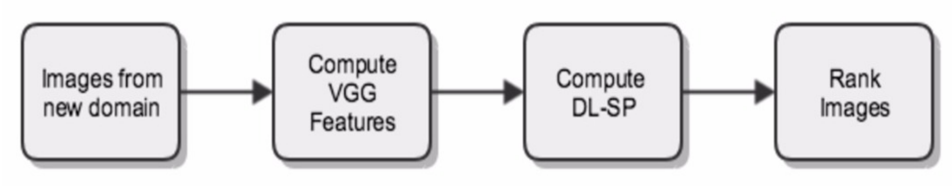


**Figure 8: Framework for cross platform image grading.**


**Figure 9: SP of Flipkart(Platform1) images and their corresponding similar styles in Myntra(Platform2).**


**Figure 10: Product Groups.**