

## Special Issue

# A Siamese Neural Network Application for Sales Forecasting of New Fashion Products Using Heterogeneous Data

Giuseppe Craparotta<sup>1,\*</sup>, Sébastien Thomassey<sup>2</sup>, Amedeo Biolatti<sup>3</sup><sup>1</sup> Department of Mathematics, University of Torino, Via Carlo Alberto 10, Torino 10123, Italy<sup>2</sup> ENSAIT, GEMTEX Laboratoire de Génie et Matériaux Textiles, F-59000 Lille, France<sup>3</sup> DISMA Department of Mathematical Sciences, Polytechnic University of Turin, Corso Duca degli Abruzzi, 24, 10129 Torino TO**ARTICLE INFO***Article History*

Received 29 Mar 2019

Accepted 19 Oct 2019

*Keywords*Sales forecasting  
Siamese neural networks  
Fashion products  
Fashion retail**ABSTRACT**

In the fashion market, the lack of historical sales data for new products imposes the use of methods based on Stock Keeping Unit (SKU) attributes. Recent works suggest the use of functional data analysis to assign the most accurate sales profiles to each item. An application of siamese neural networks is proposed to perform long-term sales forecasting for new products. A comparative study using benchmark models is conducted on data from a European fashion retailer. This shows that the proposed application can produce valuable item level sales forecasts.

© 2019 The Authors. Published by Atlantis Press SARL.

This is an open access article distributed under the CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>).**1. INTRODUCTION**

For fashion retailers, sales forecasting is crucial for planning, inventory and retail management. Sub-optimal management leads to unbalanced stock levels among stores, lost sales or excessive devaluation of the goods at the end of the season [1]. Retail sales modeling is challenging because the data are typically affected by a lot of exogenous variables: they suffer from sudden changes of customers' preferences and stores' strategy, they are affected by seasonality effects, global and local trends, weather, competition and other factors. Fragmentation is another typical feature of retail data, as one item can be sold in a number of variants, sizes and colors. Furthermore, pricing strategies strongly affect sales. All these mentioned factors interact in a complex way; some can be modified by the retailer, others cannot, and some are even non-predictable [2]. Typically, retailers aggregate their products according to a hierarchical classification providing item category, subgroups, model, color, size and sometimes fabric. Usually, some of those characteristics are codified and recorded in the company information systems.

For new products, historical sales data is not available. However, the characteristics of new items, their description and images are usually available. Presently, technology, social media and web data allow more and also new types of data to be collected. For example, images of the garments are often available. This type of data appears to be very interesting from the forecasting point of view, because the style of a fashion product is one of the most significant factors of purchase. The image of a fashion product contains a large amount

of information, even if sometimes it not easy to be described and classified. Therefore, the use of images can be valuable input for the sales forecasting of fashion products.

In existing literature, a number of studies deal with sales forecasting, based on historical sales data, at the highest levels of hierarchy, i.e., commercial categories. Managers often use three different forecasting techniques: qualitative methods [3], classical statistic techniques such as time series (Moving Average and Exponential Smoothing) or regression analysis [4]. In the last few decades, artificial intelligence techniques, such as fuzzy systems, neural network (NN) and hybrid models integrating multiple intelligent techniques, are being used more frequently in retail forecasting. Among these artificial intelligence techniques, the NN model is one of the most commonly used. It has been proved to be universal approximator and can effectively model various time series and non-time series [5]. A number of studies [6–8] also demonstrates that the NN approach outperforms the classical models due to its capacity of non-linearity and generalization.

When a forecast is made at the single Stock Keeping Unit (SKU) level for new products, no historical data is available, and as a result it is not possible to use traditional methods. For that reason, new product forecasting is often cited by many companies as one of the most difficult forecasting problems retailers face [9]. This situation is typical in the fashion industry, as most fashion collections are new products. Thus, some companies forecast sales for new products by looking for comparable historical items. This kind of match can rely on manager experience and skills, and be based on visual attributes (like style or color) or on other descriptive criteria

\* Corresponding author. Email: [giuseppe@evopricing.com](mailto:giuseppe@evopricing.com)

(Figure 1). However, it is hard to identify which past products are similar due to large product variety and frequent product changes. Furthermore, this approach is based on descriptive criteria matches, although there is no guarantee that it corresponds to a good match of sales profiles.

In [10], the authors propose a forecasting system based on clustering and classification tools which performs long-term item level forecasting adapted from the works in [11] and [12]. For new products, they studied the effects of three variables (item price, the beginning period for sales and the life span of the product, called *descriptive criteria*) on sales profile. Historical data was used to build a decision tree to carry out understandable links between descriptive criteria and sales profile prototypes. Then, the decision tree is used to associate future items with one of the defined prototypes according to the descriptive criteria. This model, called *KmDt* in the following, is characterized by the following steps:

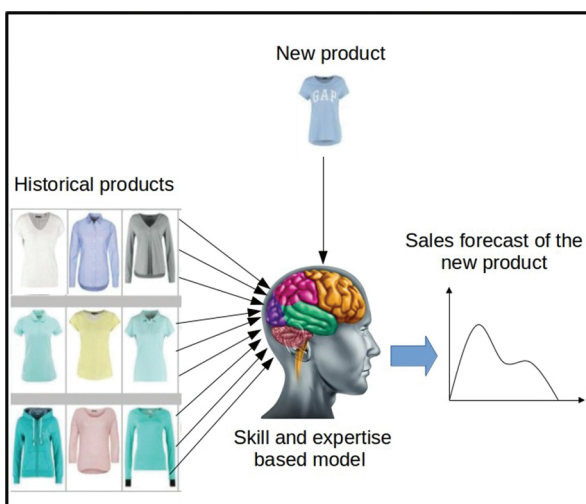
- A longitudinal clustering of sales profiles is performed
- A decision tree linking the cluster outcome to item attributes is built
- For new items, profile cluster is forecasted using the decision tree

Notice that *KmDt* model does not make use of unstructured data such as images. In [9], an implementation of diffusion models to forecast sales of a completely new-to-the-market product is proposed. The authors projected future sales with a non-linear symmetric logistic curve, using three parameters: the long-run saturation level, the inflection point of the diffusion curve, and the delay factor. Tseng [13] also exploits a diffusion model combined with fuzzy analysis, using a different reference curve (Gompertz model). In [14], a comparison between deep NNs and classical techniques such as random forests and support vector machines is performed on a SKU level. This technique achieves a similar performance compared to the random forests. The decomposition of sales in sales profiles (the shape of the sales time series) and volume (the total quantity sold) is common in these works. Actually, it is a logic

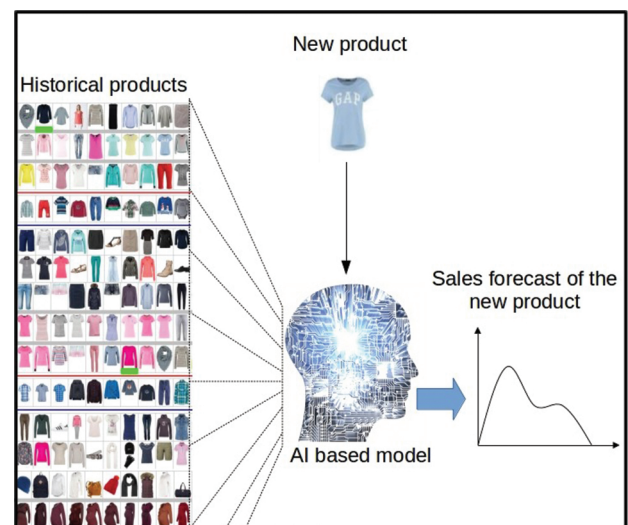
decomposition, as the sales profile is more affected by seasonality, composition and category; whereas the total quantity sold is mostly affected by distribution, stock levels, quality and trends. Due to the variety of products in a collection, the company classification system usually groups items with same seasonality: however, items in the same group may have very different potential volumes. A *naive* model to estimate the sales profile forecast can be based on a simple average of historical items with the same commercial category of the new item. This model is commonly used by managers in fashion companies as a baseline. This model will be denoted by *CategoryProfile* name in the following.

From mentioned research and the new opportunities offered by the image processing with artificial intelligent techniques, we propose in this paper a novel approach to provide SKU sales profile forecast. First, a distance between sales patterns of two historical items is defined; then, a siamese neural network (SNN) is used for learning the relation between the item characteristics and the *sales distance*. Thus, the images of the items are used as descriptors, enriching the set of descriptive criteria. Finally, it is possible to predict the sales distance between new products, with their descriptive criteria including their image, and the historical products. Historical items for which the predicted sales distance is minimum are then used as references for sales forecast. The method is inspired by the managerial process, but is completely automatic. It can manage a large number of items and attributes and avoid misleading matches (Figure 2).

We built a process using both classical descriptive criteria and unstructured data, such as images, at the same time. This allows the retailer to exploit all of the available information on the products, even from small collections. The proposed process consists of a number of preprocessing steps, a learning phase, and forecasting steps. Existing models are used to perform each step of the process; their combination and the overall process represent the contribution of this paper to the state of the art, as limited work dealt with usage of past sales, images and attributes to provide sales forecast. This approach addresses multiple issues of long-term item level forecast:



**Figure 1** | Skill based model for long-term sales forecasting of new fashion products.



**Figure 2** | AI based model for long-term sales forecasting of new fashion products.

- It allows to use images to enrich the analysis and distinguish items with similar attributes but different visual design, in order to overcome also poor annotation of attributes
- It can work well even with small fashion collections, as the maximum number of observations in the learning step is proportional to the square of the number of items
- It allows to learn complex interactions between item features, due to the peculiar learning of NNs. This is necessary as it comes out that visual attributes may be poor explanatory variables at a first glance, but their interaction with other attributes may be valuable
- It requires minimum calibration effort, as the process is almost completely data-driven. In particular, the process does not depend on any function fitting

The rest of this paper is organized as follows: Section 2 introduces the machine learning tools exploited for image processing and similarity learning: Convolutional and SNNs. The details of the proposed application are detailed in Section 3. We describe also possible similarity measures. In Section 4, a case study is presented. We applied our method to data from a European fashion retailer. We conclude in Section 5.

## 2. CONVOLUTIONAL AND SNNs

Among AI techniques, NNs emerge as the most popular and suitable to deal with unstructured data, such as text [15], conversation [16] or image [17]. However, NNs require a careful selection of input variables and network parameters such as the learning rate, the number of hidden layers, and the number of nodes in each layer in order to achieve satisfactory results [18]. It is also important to reduce dimensionality to improve learning efficiency. On the other hand, deep learning automatically extracts features from data and requires minimal human intervention during feature selection [19,20].

### 2.1. Convolutional Neural Networks

Convolutional Neural Network (CNN) is a type of feed-forward artificial NN. CNNs make the explicit assumption that the inputs are complex, which allows certain properties to be encoded into the architecture, making the forward function more efficient to implement and reducing the amount of parameters in the network. They have wide applications in image and video recognition [21], recommendation systems and natural language processing [22]. These systems can operate at the pixel level and learn both low-level features and high-level representations in an integrated manner, while they are more robust than classical techniques to geometric distortions of the input image. In order to extract features from input data, CNNs are typically composed of a combination of several layers; the convolutional layer, the pooling layer and the fully connected layer, as shown in Figure 3.

Recently, CNNs have been largely used in the fashion industry, mainly in garment recognition [23] and recommendation systems [24,25], due to their capability to learn the features of cloth representation and to easily evaluate image similarity. Few works use

CNNs to improve sales forecasting. [26] demonstrated that forecasting of long-term fashion trends could greatly benefit from visual analysis. However, this work is focused on total yearly sales of a style of garment, and therefore does not consider the sales profile of a particular SKU.

### 2.2. Siamese Neural Networks

A SNN consists of two twin CNNs which accept distinct inputs and are joined by an energy function at the top. This function computes a distance between the highest-level feature representation on each side, as shown in Figure 4. The parameters between the twin networks are tied, leading to a simpler model. SNNs ensure the consistency of predictions and symmetry, due to the identical inner working of the networks.

SNNs were first introduced to solve image matching problems [27], in particular the signature verification. Later, [28] used the architecture with discriminative loss function for face verification. Recently, these networks are used extensively to enhance the quality of visual search, for instance face recognition at Facebook [29], or even to learn the similarity metric for question–question pairs by leveraging the question–answer pairs available in Community Question Answering (cQA) forums [30].

In view of the above works, it obviously appears that

- Image analysis of fashion products should provide a valuable improvement for fashion sales forecasting

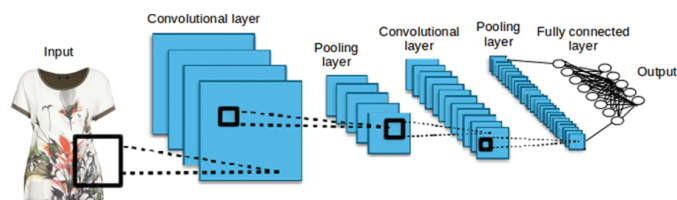


Figure 3 | Convolutional Neural Network architecture.

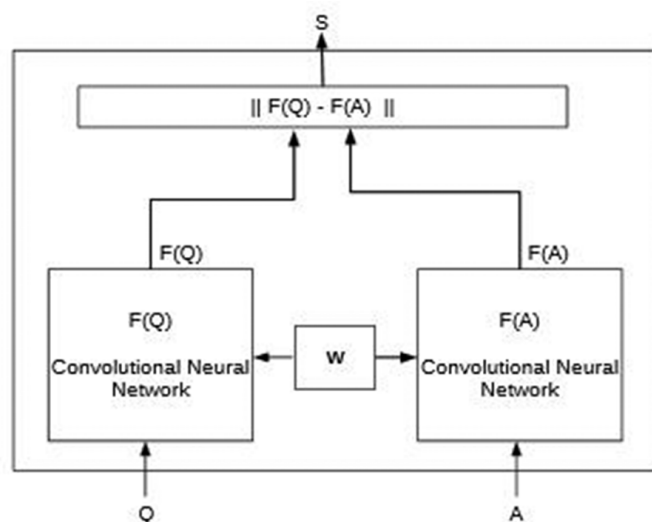


Figure 4 | Siamese Neural Network architecture.

- Convolutional and SNNs emerge as the most suitable techniques to implement this image processing in sales forecasting

Thus, a system that aims to enhance the sales forecasting of new fashion products from images is proposed.

### 3. PROPOSED PROCESS

This study presents a sales forecasting process based on SNNs for new fashion products. The method consists of a learning step in which the model is calibrated, and a test step where performances of the model are evaluated and compared to benchmark models. A flowchart of the process is presented in Figure 5, whereas the steps are detailed in the following. The whole method takes into account sales data, images and attributes of historical fashion products as input data, and provides the sales profiles of new products.

#### 3.1. Learning Phase

The learning phase allows the parameters of the model to be optimized and the most suitable similarity metric to be defined. The model is calibrated using sales of a part of the historical products (training set) to predict the sales profile of another set of historical products (validation set). Specifically, a k-fold schema is used to calibrate the prediction.

The learning phase is composed of 4 stages as follows:

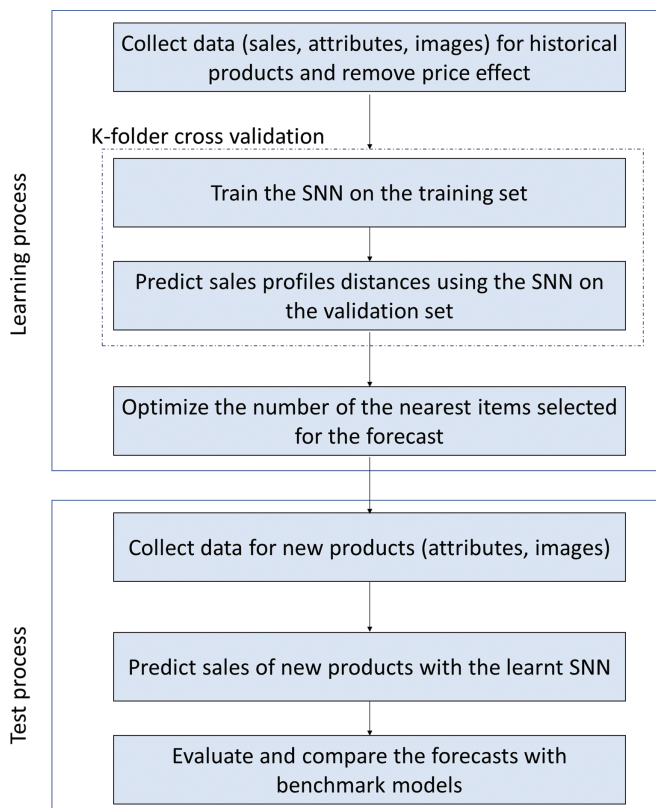


Figure 5 | Flowchart of the proposed methodology.

1. Data collection and pretreatment
2. Training of the SNN to model relationship between sales profile distances and feature distances on the training set of historical products
3. Prediction of sales profile similarity on the validation set of historical data
4. Optimization of the number of the nearest items selected for the forecast

Notice that the stages 2 and 3 are reiterated inside the k-fold cross validation process.

##### 3.1.1. Data collection and pretreatment

Sales of a single season are collected and grouped by week. For each item  $i$  and week  $w$ ; total quantity sold  $Q_{i,w}$ , average paid price and initial price are computed. Then, the quantity is normalized to obtain the sales profile  $\tilde{Q}_{i,w}$  in such a way that  $\forall i, \sum_w \tilde{Q}_{i,w} = 1$ . These profiles are strongly affected by promotional activity that can change along with the season. Promotional activity can push sales of some products and as a result, change the *baseline* sales profile. When forecasting the sales of a new product, we are interested in the sales profile at *full price* or, in other words, *price-normalized sales*. In order to remove impact of price initiatives on sales, we clean the sales time series using the elasticity of the demand with respect to price [31–33] (See Appendix for further details). Figure 6 shows an example of sales profiles before and after price normalization.

The direct effect of the price normalization is that sales profiles are less affected by price discount. Furthermore, this step indirectly guarantees that the distances between sales profile are slightly affected by punctual price initiatives.

##### 3.1.2. K-fold cross validation

For a given item  $i$ ,  $\tilde{Q}_{i,w}$ , the quantity sold at week  $w$ , is a real function of the discrete variable  $w$ . We can thus define the distance *in sales* between two items  $i_1$  and  $i_2$  using any distance function  $d_{i_1,i_2}$ . We chose to define  $d$  as the root mean square deviation between the two functions  $\tilde{Q}_{i_1}$  and  $\tilde{Q}_{i_2}$ , in order to account for the whole sales history:

$$d_{i_1,i_2} = \|\tilde{Q}_{i_1} - \tilde{Q}_{i_2}\|_2 = \sqrt{\sum_w (\tilde{Q}_{i_1,w} - \tilde{Q}_{i_2,w})^2}$$

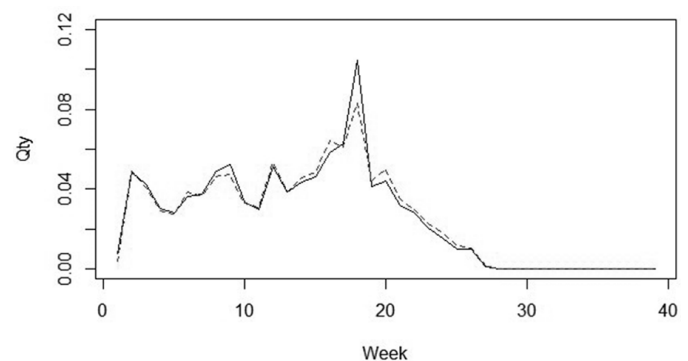


Figure 6 | Example of sales profiles before and after price normalization.



The aim of the SNN is to model the relation between distance  $d_{i_1,i_2}$  and the similarity in the feature space (input data). As for signature verification, the role of the SNN model is to learn which features of a garment have to be similar in order to guarantee a similarity in sales profile.

Initially, item features are fully defined; categorical features are transformed in dummy variables and the set of numeric and dummy features is then joined with image features. Therefore, the standard SNN architecture is slightly modified, as shown in Figure 7, to integrate structured data as tags and attributes with the information derived from item images. This modification is needed to join all available data (tags and images) which constitute the complete DNA of the garment, as defined in Bracher et al. [34]. To be specific, input data are pairs of historical items composed by the attributes and the image of the items. Every image is then codified through a CNN. Finally, the attributes and CNN outputs are joined together and the distance between the features of the pair are compared to the sales distances.

In order to train the model using historical data, the set of historical items  $I$  is partitioned in  $k$  subsets  $I_1, I_2, \dots, I_k$ , and  $k$  models are built. To build the  $f^{th}$ , with  $f = 1, \dots, k$ , training dataset all products  $i_1$  and  $i_2$  in  $I \setminus I_f$  are pairwise matched, features of each item in each pair are collected, and distances  $d_{i_1,i_2}$  are computed. To build the  $f^{th}$  validation dataset, all items  $i_1$  in  $I_f$  are matched with all items  $i_2$  of the complementary set  $I \setminus I_f$ . Features of each item in each pair are collected, and distances  $d_{i_1,i_2}$  are computed.

It is possible to obtain many variations of SNN by changing the structure of CNN and the energy function. We tested two possibilities for the CNN:

- The 50 layer Residual Network (ResNet50) as proposed in [35] and pre-trained on the ImageNet data set [36]
- A simpler CNN (without pre-training) constituted of three convolutions with a ReLU non-linearity interspersed with max-pool layers

The first model is by far more complex, with about 23 million parameters against 260 thousand of the second. Applying these models on 200x200 images, ResNet requires about  $6.6 \cdot 10^9$  multiply-accumulate operations against  $3.7 \cdot 10^9$  of the second one and in practice is from 3 to 4 times slower. Obviously, the higher

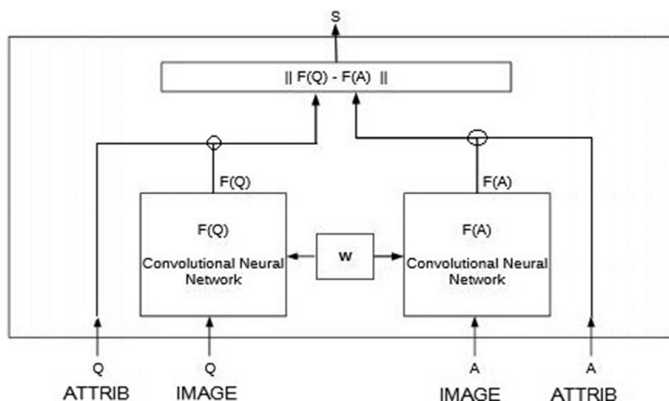


Figure 7 | The modified Siamese Neural Network (SNN) architecture.

computational burden is compensated by a higher representative power. These two models have obtained similar results on data. Consequently, we choose to only explain the methodology based on the simple CNN in the following. This is probably due to the relatively small number of examples, which do not allow the more complex model to capture high-level features. For the energy function, there are also different possibilities. The simplest approach is to directly use a summarizing function to condensate the couple of feature vectors in the prediction. Examples of suitable functions are the  $L_2$  norm of the difference of the vectors, the cosine distance or, more generally, any distance function defined on the features space. This approach has the advantage of being straightforward, requiring only to define and train the CNN, although it also has some downsides. First, it does not discriminate between the dimensions of the encoding vectors. Therefore, it cannot recognize correlations between them. Secondly, it is not clear how we could add more information than the images, namely the attributes, if available. This latter point is particularly important in our case, since different information is often available (category, price, etc.) that could ease the learning process or even lead to better results. A more general approach could be to learn the energy function itself with an additional NN, that can be interpreted as a learnable distance, with the concatenation or the difference of the features returned by the CNN as an input [37]. Using the absolute value of the difference of the features in input of this NN ensures the symmetry of the SNN. In this work we used this approach based on a learnable distance function, using the concatenation of the features as input. The resulting model, composed of two CNNs with shared weights and a fully connected NN, can be seen as a NN itself and therefore can be trained with standard approaches, such as gradient descent via back-propagation. At the end of this cross validation step,  $f$  SNNs have been trained and a prediction of the distance  $\tilde{d}_{i,j}$  between for all the pairs of items  $i$  and  $j$  of  $I$  is given.

### 3.1.3. Parameter optimization

From the above predicted distances  $\tilde{d}_{i,j}$  between a pair of items  $i$  and  $j$  of  $I$ , the sales profile forecast can be performed for each item in  $I$ . The sales profile forecast for an item  $i \in I_f$  is defined as the average sales profile of items  $j$  in  $I \setminus I_f$  for which the predicted distance  $\{\tilde{d}_{i,j}\}_j$  is low. Therefore, the number  $n$  of nearest profiles to consider has to be determined. We define  $L_i^n$  as the set of items having the  $n$  lowest predicted distance from the sales profile of item  $i$ . Thus, the sales profile forecast for item  $i$  is

$$\hat{Q}_i^n = \frac{\sum_{j \in L_i^n} \tilde{Q}_j}{n}$$

In the learning phase, this prediction can be compared to real normalized sales profiles to compute the total forecast error. There are several ways to measure forecast performance, including evaluating the forecast results in terms of accuracy, error, cost, efficiency, profit, and/or customer satisfaction. Common measurements (e.g., Mean Absolute Percentage Error (MAPE), Mean Absolute Deviation (MAD), and Root Mean Squared Error (RMSE)) have all been used by companies to assess the forecasting function. In this work, the use of RMSE is suggested:

$$\varepsilon^n = \sqrt{\frac{1}{\text{card}(I)} \sum_{i \in I} (\hat{Q}_i^n - \tilde{Q}_i)^2}$$

The choice of  $n$  is a crucial step for the forecast accuracy. On one hand, if  $n$  is too low, prediction can be affected by some peculiar nearest profiles. On the other hand, if  $n$  is too high, the model would take into account items that are not near the predicted sales profiles.

Thus, we calibrate the model by identifying the number  $n$  minimizing total forecast error  $\epsilon^n$ .

### 3.2. Test Phase

Once the model is fully defined, our procedure is tested on a set of new items of a new season.

The test phase consists of

- Data collection for new products
- Prediction of sales profiles of new products
- Evaluation of the forecast errors and comparison with benchmark models

In order to evaluate performances of the application proposed in this work, forecasts are also made using two other methods dealing with the same forecast instance (long-term, item level, new product forecast), *CategoryProfile* and *KmDt* methods, already defined in the related work section.

## 4. EMPIRICAL STUDY

We applied the proposed method to real data from a European fashion retailer. We used historical data (sales, attributes and images) of womens tops from the 2015 and 2016 summer collections, and tested the methodology using sales from the 2017 summer collection, attributes and images. Thus, the dataset is built using 360 products from the historical collections and 88 products from the new collection. For each product, the following attributes are taken into account: initial price, life span, beginning period, subcategory, fabric and type of drop.

### 4.1. Model Learning

The learning process described in Section 3.1 is implemented on the 360 historical items.

The SNN used in this case is composed of two CNNs with three convolutional layers alternating with max-pool layers. The two output feature vectors relative to the two input images are chained together. Additional input vectors, composed of attributes are concatenated, if available. The resulting vector is then fed to a dense NN of three layers that returns the desired output. The non-linearity is ensured by rectified linear units (ReLU) for both the convolutional and the dense layers. Figure 8 shows the RMSE during the learning phase on the training and validation data. We used approximately 40.000 iterations of gradient descent, using the well known Adam method [38]. To avoid overfitting, dropout is used during the training, after the second layer of the dense NN with a drop rate of 0.5. This technique, first introduced in [39], randomly turns off some elements of a layer independently at each iteration to avoid co-adaptations of

the features that can cause overfitting. The stagnation of the validation error indicates that the learning process is finished.

The learning process of the SNN enables the distances between sale profiles to be predicted for each product. Figure 9 proposes a synthetic view of the real and the predicted distances between profiles of one product  $i \in I_f$  and other products in  $I \setminus I_f$ . The point cloud is clearly diagonally oriented meaning that the SNN obtained after the learning is able to provide an accurate forecast of the distances.

From these predicted distances, the selection of the best number  $n$  of nearest products is then performed. Following the process described in Section 3.1, the profile forecasts are computed for  $n = 4$  to 60. As expected,  $n$  is a sensitive parameter of our methodology: a too low or too high value of  $n$  leads to an increase of the RMSE. Figure 10 shows that the best average RMSE is found for  $n = 20$ .

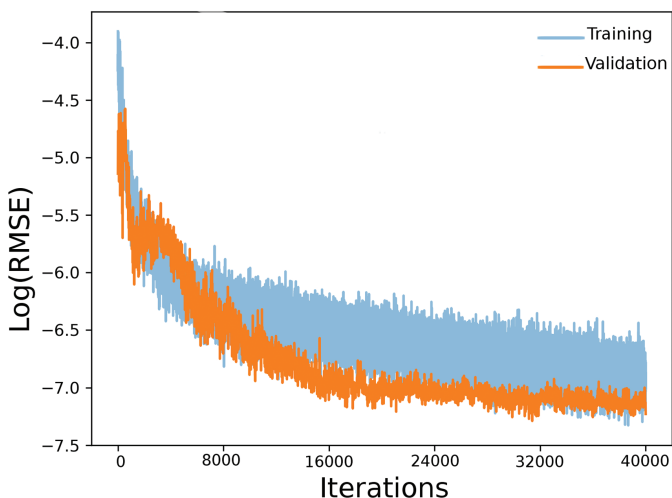


Figure 8 | Root mean squared error (RMSE) during the learning phase of the Siamese Neural Network (SNN).

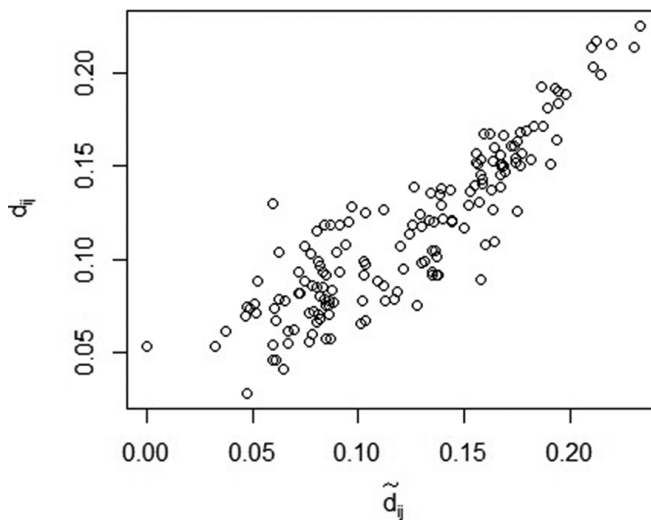


Figure 9 | Comparison of real  $(d_{i,j})$  and predicted  $(\tilde{d}_{i,j})$  profile distances between one item  $i \in I_f$  and other items  $j \in I \setminus I_f$ .

## 4.2. Profile Forecast of New Products and Result Comparisons

The trained SNN with  $n = 20$  nearest products, called SNN20, is implemented for profile forecasting of the 88 new products. The profile forecasting process of a new product is illustrated in Figure 11. To facilitate reading, the number of nearest products given by the SNN is deliberately limited to  $n = 3$  in this example (instead of 20 in the proposed model).

Sales profile forecasting is also performed using the benchmark models described in Section 3.2: *CategoryProfile* and *KmDt*. The average RMSE obtained in the test data are presented in Table 1.

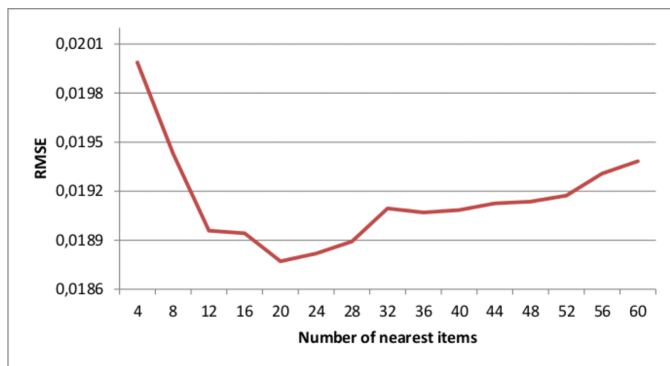


Figure 10 | Root mean squared error (RMSE) for normalized sales profile with respect to the number  $n$  of nearest predicted items.

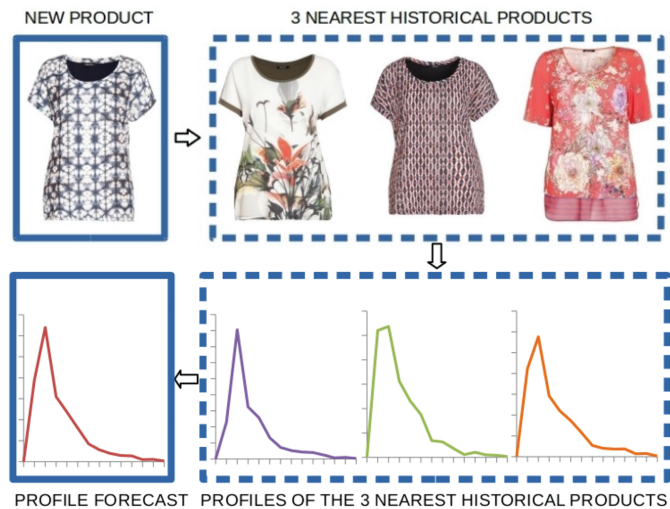


Figure 11 | Example of profile forecasting process of a new product with the 3 nearest products.

Table 1 | Comparison of mean RMSE on test data.

	Mean RMSE	% of RMSE Improvement with SNN20
<i>Profile Category</i>	0.0234	19.6%
<i>KmDt</i>	0.0199	5.5%
SNN4	0.0200	6%
SNN60	0.0194	3%
SNN20	0.0188	-

RMSE, root mean squared error.

Considering these overall results, our model SNN20 (with  $n = 20$ ) outperforms the two models used for comparison. It also can be noted that a lower or a larger value of  $n$  decreases the performance of the proposed model in the test data (see RMSE of the SNN4 and SNN60 models). However, a more specific analysis of the results demonstrates that a few issues arise on some products. Figure 12 proposes a boxplot of the percentage of RMSE improvement of our model compared with competitors. If very few forecasts are better for the *ProfileCategory* model, almost one quartile of the new products obtains a better forecast with the *KmDt* model.

Figure 13 illustrates the same observation with a distribution of the RMSE per product and gives a more detailed view of the failures of the SNN20 model.

From the analysis of this figure, we extract the two worst forecasts of the SNN20 model compared to the *CategoryProfile* model, which is the less accurate model, according to Table 1. The sales profile curves are compared in Figure 14. Two main explanations can be formulated to motivate the failure of the proposed model:

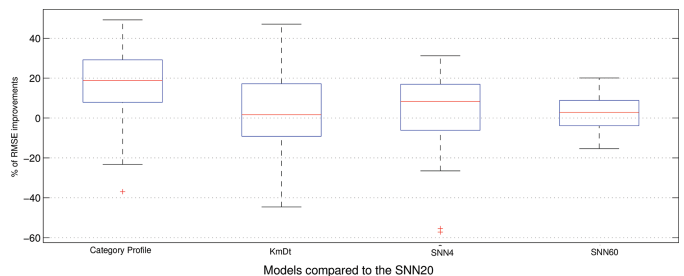


Figure 12 | Boxplot of root mean squared error (RMSE) improvement between the SNN20 model and the other models.

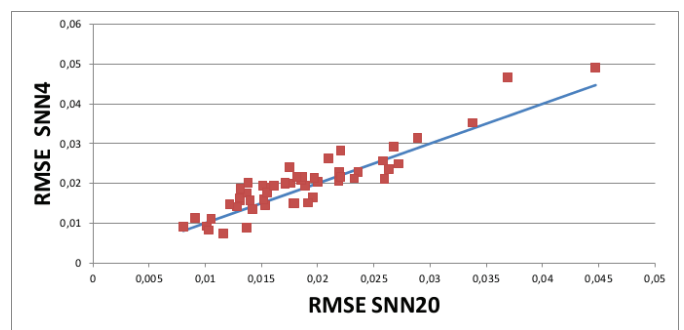
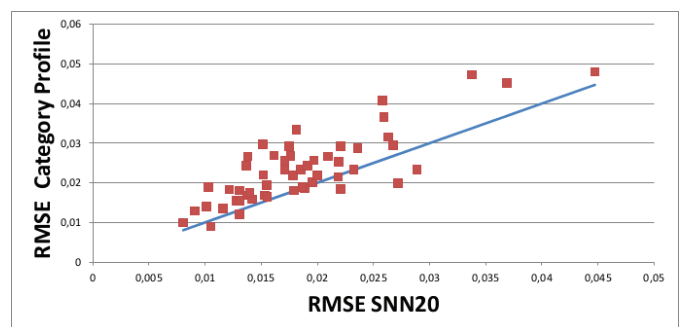


Figure 13 | Comparison of the root mean squared error (RMSE) per product between the SNN20 model and the other models.

- For the products A and B, actual sales are very fluctuating and erratic. This can be explained, for instance, by the influence of unexpected exogenous factors. In a such situation, the sales are very difficult to predict and it is consistent that the simplest and most averaged model gives the best results.
- For products C and D, the main problem arises from the beginning period of the sales. This issue cannot exist in real situations since the beginning period of the sales are mastered by the company. As a result, the manager can synchronize the forecast to the right date. This issue will be subject to future works. It can be noticed that the forecasts provided by the SNN20 model are still acceptable compared to the naive model.

Similarly, we extract the 4 best forecasts of the SNN20 model compared to the *KmDt* model which is the most accurate competitor model according to Table 1. For these four products, E, F, G and H, the accuracy of the SNN20 model is very satisfactory for a long-term sales forecasting of new product as illustrated in Figure 15.

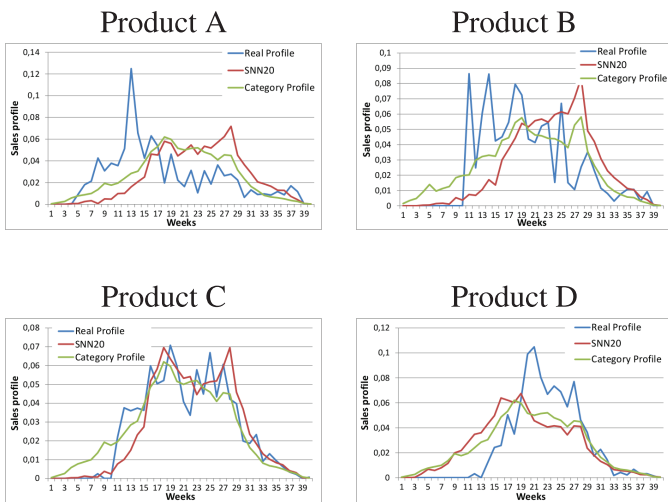


Figure 14 | The 4 worst forecasts of the SNN20 model compared to the model.

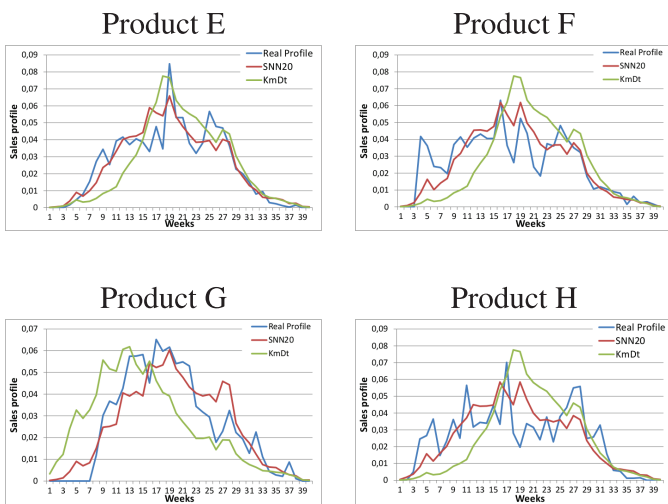


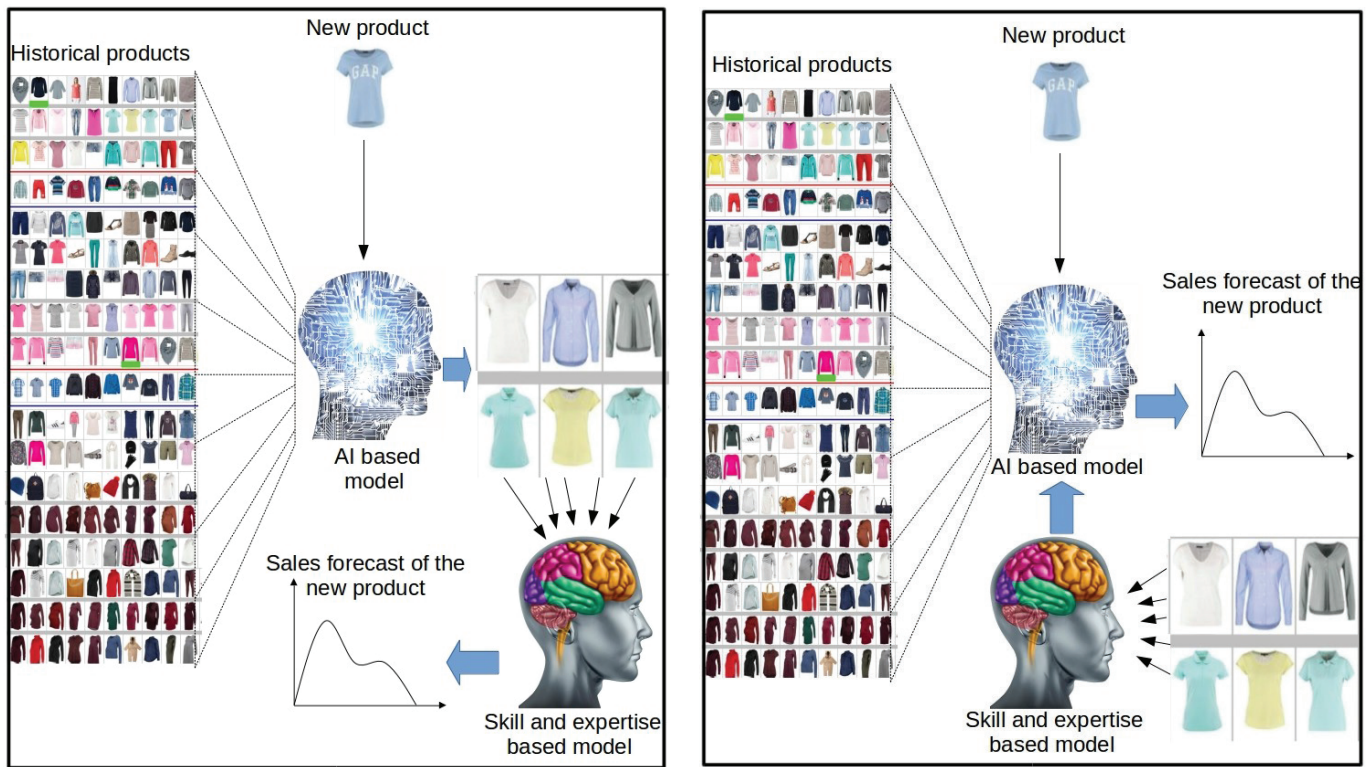
Figure 15 | The 4 best forecasts of the SNN20 model compared to the model.

## 5. CONCLUSION AND PERSPECTIVES

A novel process was developed to forecast sales profiles for new fashion products. This process, consisting of a number of preprocessing steps, a learning phase and forecast steps using existing models, allows to build new product long-term item level forecast. In addition, it addresses a number of issues found in other methods such as the usage of multiple types of data. This model is based on a SNN which allows for the use of numerical and categorical input variables, as well as unstructured data such as images of the product. The main concept of the proposed process is based on the best practices, used in fashion companies, which consists of comparing the design, the style, the visual appearance and other technical attributes of new products to those of historical ones in order to perform sales forecasting. This method enables the most experienced and trained operators to generate quite satisfactory forecasts, but it is often limited by the number of products and attributes which are taken into account. This task can also be very time consuming for managers who have a considerable amount of other issues to deal with. Thus, the proposed method based on artificial intelligence techniques, aims to enhance this practice with a wider and deeper analysis of historical products. The proposed process is applied to real data from a European fashion retailer. Overall, the performance obtained with our model is better than the two models used for comparison: the *CategoryProfile* (or naive) model, which is mostly used by companies for such forecasts, and the *KmDt* model which is based on clustering and decision tree techniques. These results demonstrate that images of the products are valuable inputs for sales forecasting. In addition, contrary to the two competing models considered for comparison, the developed method is not affected by classification errors. However, if performances of the proposed model are generally better, a more specific analysis of the results shows that the model fails to overcome the naive model for some products. Thus, different improvements can be envisaged in future works:

- The beginning period of sales has to be integrated in a better way in the model. This date is considered as an attribute in the current model but it should be taken into account directly in the final forecast.
- The training of the SNN can be consolidated with a more suitable and/or a larger database. The current model has been learned from scratch on a database of tops and T-shirts for women. A future investigation could consist in performing a pre-training of the model on a wider database including various heterogeneous categories, and a fine tuning on a specific database composed of the target category of garments. To balance the trade-off between the accuracy and the computational cost pruning techniques, as described in [40], could be used to reduce the effective size of the network. Alternatively, the features of the reference items could be pre-computed to speed up the inference.
- The number of nearest products to achieve the final forecast is a crucial parameter. The optimization of this number could be further examined, for instance, by customizing this number for each product. The ranking of the historical products is also an interesting aspect which could disclose the most distinctive products. Furthermore, for each product, a confidence index





**Figure 16** Hybrid models for long-term sales forecasting of new fashion products.

could be computed from these rankings to indicate if the sales forecast is reliable.

- The proposed method is fully automated and exclusively relies on artificial intelligence. This kind of model is successfully implemented for digital business but could be not accepted in fashion companies where subjectivity and creativity are the core business. In such a situation, a hybrid approach should be more suitable. Different methodologies can be considered, for instance, the models presented in Figure 16. The AI based model can help the operator by providing a set of potential products to take into account. Then the final forecast is produced with the skills and knowledge of the operator from this set (Figure 16 left). The human skills and knowledge can be also integrated into the AI based model to enhance the efficiency and direct the final result (Figure 16 right).

To conclude, long-term sales forecasting of new fashion products is a very challenging problem. As demonstrated in this work, the last advances in artificial intelligence and big data management offer new opportunities to deal with this fascinating issue. There is no doubt that further improvements will emerge in the near future.

## AUTHORS' CONTRIBUTIONS

Giuseppe Craparotta contributed to state of the art and model design, data preprocessing, benchmark and model calibration. Sébastien Thomassey contributed to state of the art and model design, benchmark and results analysis. Amedeo Biolatti contributed to design, implement, test and verify the proposed model.

## Funding Statement

This work was supported by ENSAIT (École nationale supérieure des arts et industries textiles).

## ACKNOWLEDGMENTS

The authors gratefully acknowledge Evo Pricing and ENSAIT for promoting and collaborating to this research.

## REFERENCES

- [1] M.E. Nenni, L. Giustiniano, L. Pirollo, Demand forecasting in the fashion industry: a review, *Int. J. Eng. Bus. Manage.* 5 (2013).
- [2] S. Thomassey, Sales forecasts in clothing industry: the key success factor of the supply chain management, *Int. J. Prod. Econ.* 128 (2010), 470–483.
- [3] N.R. Sanders, K.B. Manrodt, Forecasting practices in U.S. corporations: survey results, *Interfaces.* 24 (1994), 92–100.
- [4] I. Alon, M. Qi, R.J. Sadowski, Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods, *J. Retailing Consum. Serv.* 8 (2001), 147–156.
- [5] G. Cybenko, Approximations by superpositions of sigmoidal functions, *Math. Control Signals Syst.* 2 (1989), 303–314.
- [6] Y. Yu, T. Choi, C. Hui, An intelligent fast sales forecasting model for fashion products, *Expert Syst. Appl.* 38 (2011), 7373–7379.
- [7] K.F. Au, T.M. Choi, Fashion retail forecasting by evolutionary neural networks, *Int. J. Prod. Econ.* 114 (2008), pp. 615–630.
- [8] Z.L. Sun, T.M. Choi, K.F. Au, Y. Yu, Sales forecasting using extreme learning machine with applications in fashion retailing, *Decis. Support Syst.* 46 (2008), 411–419.

- [9] J.T. Mentzer, M.A. Moon, *Sales Forecasting Management: a Demand Management Approach*, Sage Publications, 2005.
- [10] S. Thomasse, A. Fiordaliso, A hybrid sales forecasting system based on clustering and decision trees, *Decis. Support Syst.* 42 (2006), 408–421.
- [11] C. Ching-Chin, A.I.K. Ieng, W. Ling-Ling, K. Ling-Chieh, Designing a decision-support system for new product sales forecasting, *Expert Syst. Appl.* 37 (2010), 1654–1665.
- [12] S. Thomasse, M. Happiette, J.M. Castelain, A global forecasting support system adapted to textile distribution, *Int. J. Prod. Econ.* 96 (2005), 81–95.
- [13] F.M. Tseng, Quadratic interval innovation diffusion models for new product sales forecasting, *J. Oper. Res. Soc.* 59 (2008), 1120–1127.
- [14] A.L.D. Loureiro, V.L. Miguisa, L.F.M. da Silva, Exploring the use of deep neural networks for sales forecasting in fashion retail, *Decis. Support Syst. J.* 114 (2018), 81–93.
- [15] R. Moraes, J.F. Valiati, W.P.G. Neto, Document-level sentiment classification: an empirical comparison between SVM and ANN, *Expert Syst. Appl.* 40 (2013), 621–633.
- [16] Z. Yan, N. Duan, J. Bao, P. Chen, M. Zhou, Z. Li, Response selection from unstructured documents for human-computer conversation systems, *Knowl. Based Syst.* 142 (2018), 149–159.
- [17] V. Andrearczyk, P.F. Whelan, Convolutional neural network on three orthogonal planes for dynamic texture classification, *Pattern Recognit.* 76 (2018), 36–49.
- [18] A.J. Hussain, A. Knowles, P.J.G. Lisboa, W. El-Deredy, Financial time series prediction using polynomial pipelined neural networks, *Expert Syst. Appl.* 35 (2008), 1186–1199.
- [19] C. Affonso, A.L.D. Rossi, F.H.A. Vieira, A.C.P.D.L.F. Carvalho, Deep learning for biological image classification, *Expert Syst. Appl.* 85 (2017), 114–122.
- [20] E. Chong, C. Han, F.C. Park, Deep learning networks for stock market analysis and prediction: methodology, data representations, and case studies, *Expert Syst. Appl.* 83 (2017), 187–205.
- [21] S. Lawrence, C.L. Giles, A.C. Tsoi, A.D. Back, Face recognition: a convolutional neural-network approach, *IEEE Trans. Neural Netw.* 8 (1997), 98–113.
- [22] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, T. Chen, Recent advances in convolutional neural networks, *Pattern Recognit.* 77 (2018), 354–377.
- [23] Z. Li, Y. Sun, F. Wang, Q. Liu, Convolutional neural networks for clothes categories, in *CCF Chinese Conference on Computer Vision*, Springer, Berlin, Heidelberg, 2015, 120–129.
- [24] J. McAuley, C. Targett, Q. Shi, A. Van Den Hengel, Image-based recommendations on styles and substitutes, in *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Santiago, Chile, 2015.
- [25] Y. Qian, P. Giaccone, M. Sasdelli, E. Vasquez, B. Sengupta, Algorithmic clothing: hybrid recommendation, from street-style-to-shop, in *Proceedings of KDD*, Canada, 2017. arXiv:1705.09451.
- [26] Z. Al-Halah, R. Stiefelwagen, K. Grauman, Fashion forward: forecasting visual style in fashion, in *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, 2017, pp. 388–397.
- [27] J. Bromley, I. Guyon, Y. LeCun, E. Sackinger, R. Shah, Signature verification using a “siamese” time delay neural network, *Adv. Neural Inf. Process. Syst.* 7 (1993), 737–744.
- [28] S. Chopra, R. Hadsell, Y. LeCun, Learning a similarity metric discriminatively, with application to face verification, in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, San Diego, 2005, pp. 539–546.
- [29] Y. Taigman, M. Yang, M. Ranzato, L. Wolf, DeepFace: closing the gap to human-level performance in face verification, in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, 2014, pp. 1701–1708.
- [30] A. Das, H. Yenala, M. Chinnakotla, M. Shrivastava, Together we stand: siamese networks for similar question retrieval, in *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, Berlin, 2016, pp. 378–387.
- [31] I. Png, *Managerial Economics*, Blackwell Publishers, 1998. ISBN 1557869278.
- [32] M. Parkin, M. Powell, Melanie, K. Matthews, *Economics*, Addison-Wesley, 2008. ISBN 0132041227.
- [33] A. Gillespie, *Foundations of Economics*, second ed., Oxford University Press, 2011. ISBN 0199586543.
- [34] C. Bracher, S. Heinz, R. Vollgraf, Fashion DNA: merging content and sales data for recommendation and article mapping, in *Machine Learning Meets Fashion Workshop, KDD 2016 Conference*, San Francisco, 2016. arXiv:1609.02489.
- [35] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 2016, pp. 770–778.
- [36] O. Russakovsky, J. Deng, H. Su, *et al.*, ImageNet large scale visual recognition challenge, *Int. J. Comput. Vision.* 115 (2015), 211–252.
- [37] S. Zagoruyko, N. Komodakis, Learning to compare image patches via convolutional neural networks, in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 2015, pp. 4353–4361.
- [38] D.P. Kingma, J. Ba, Adam: a method for stochastic optimization, in *3rd International Conference for Learning Representations*, San Diego, 2015. arXiv:1412.6980.
- [39] G.E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskeve, R.R. Salakhutdinov, Improving neural networks by preventing co-adaptation of feature detectors, arXiv preprint arXiv:1207.0580, 2012.
- [40] P. Molchanov, S. Tyree, T. Karras, T. Aila, J. Kautz, Pruning convolutional neural networks for resource efficient inference, in *Proceedings of the International Conference on Learning Representation (ICLR 2017)*, Toulon, France, 2017. arXiv:1611.06440.