変動するグローバル電子デバイス市場における在庫最適化に向けた時間的センチメント・トピック分析技術

Demand forecast contextual explainability through topic and temporal Sentiment analysis

コルコスサヒム* ポールデブディープ** ガヤトリサラナタン** 三浦 康史*** 多鹿 陽介***
Sahim Kourkouss Paul Debdeep Saranathan Gayathri Koji Miura Yosuke Tajika

要 旨

近年のパンデミックや半導体不足により、グローバル電子デバイス市場は激しく変動している。これにより、顧客の発注見込みが大きくぶれ、過剰在庫や販売機会損失を引き起こしている。これに対し、我々はマーケットの最前線で営業活動を行っているグローバル 9 地域 100 名以上の営業の景況感とその理由を示す自然言語によるコメントを収集・活用することにより、当社グローバル電子デバイスの需要を予測する AI 技術を開発した。当社営業の自然言語によるコメントは当社ドメイン知識を含む将来の景況が記載されるとともに、これらの因果関係やニュアンスなど数値データにはない情報が含まれている。筆者らはこれらのデータから予測に有用な情報として、将来の景況の良し悪しを表現する時間的センチメント情報、および、それらが関係するマーケットを特定するトピック情報を抽出・分類する学習モデル開発し、予測モデルへの適用を行った。これにより、当社商品に適合した予測の妥当性説明を実現する。また、予測説明情報を営業担当者に予測結果と共に提供し、その妥当性・信頼性において高評価を得た。

Abstract

While the digitalization increase in manufacturing provides a growing amount of data that can describe assets and operations, supply chains as well as customer purchase patterns are becoming increasingly complex, and therefore harder to understand or predict. Recent Artificial intelligence models can resolve such complexity by using their ability to sort and analyze massive amounts of data and providing accurate forecasts. However, one of the biggest challenges for business executives today is to quantitatively grasp the demand volatility to make informed decisions. High accuracy alone is not enough to realize such responsible decision making, but it is also required to have an in-depth knowledge of the model's behavior to improve confidence as well as create more realistic expectations. In previous works [1] we have developed models for accurate forecasts regarding stable and irregular demand periods. In this paper we propose a novel method that combines topic modeling and temporal sentiment analytics to provide a better understanding of our AI forecasting models. Our aim is to help users make informed decisions and avoid errors that could result in significant losses. Such explanations reinforce our AI models, increase confidence in the system, and help identify errors and performance problems. As an evaluation method, we compare our proposed method with conventional cases where the forecast is generated through relevant macroeconomic indices. We have presented our findings to internal sales department and have received highly rated feedback for this concept.

1 Introduction

Web news, search engines, and posts on social media, new data is constantly being produced in society today, and this has led to data science becoming a booming field in which all these data are analyzed to extract information and to generate new values.

In the manufacturing industry, it is a fact that amid drastic changes in the market structure, rapid and accurate grasp of demand trends, realizing flexible production planning to cope with fluctuations and optimal supply to customers is a vital requirement. For this purpose, there is an urgent need to collect big data such as sales and manufacturing data, and to effectively apply data-driven approaches based on Artificial intelligence technology to said sales and manufacturing activities.

In previous works [1], we have developed intelligent models capable of providing accurate forecasts regarding stable and irregular demand periods and have proven that AI-based quantitive approaches can resolve complex forecast problems by using their ability to sort and analyze massive amounts of data.

In this paper we go a step further and propose a novel method that provides better understanding and explicability of such AI forecasts, by combining data mining techniques such as topic modeling and explainable machine learning through temporal sentiment analytics.

* パナシンポインダストリー 技術本部 グローバル戦略部
Global Strategy Department, Engineering Division, Panasonic Industry
** パナソニックデバイスシンガポール シンガポール技術開発センター
Singapore Technology Center, Panasonic Industrial Devices Singapore
*** パナシンポインダストリー 技術本部 プロセスデバイス革新センター
Process Device Innovation Center, Engineering Division, Panasonic Industry
問合せ先：miura.koji@jp.panasonic.com

2 Background

2.1 The long tail sales strategy

With the recent dramatic market structure changes in global electronics sales, the use of distributors for long-tail customers has become a key to profitable growth. Our company's electronic components sales to long-tail customers have been increasing on a yearly basis and has now become more important than ever.

[Fig.1] shows the product supply chain for sales through a distributor. The supply chain consists of two mutually influencing parts: the distributor's side and the manufacturer's side. While the distributor's part represents mainly the relationship between the distributor for long-tail customers and the long-tail customer itself, the manufacturer's side embodies the relationship between the distributor and the manufacturer (our company). The complexity of such supply chains as well as the huge number of purchase patterns and the infrequency for long tail sales makes it very challenging to produce accurate forecasts based only on conventional qualitative forecasting methods such as subjective opinions and insights.

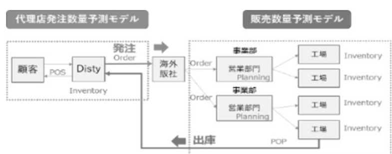

Fig. 1: Product Supply Chain and Models

2.2 Importance of explainability in demand forecast

In our previous works, we have developed a technology for predicting accurate sales volumes on a part-numbered level for long-tail customers, and we have applied it to the sales planning process in various sales departments within our company, thereby reducing planning person-hours by over 80% [1].

However, one of the biggest challenges for sales planners and business executives today is to make informed decisions by quantitatively grasping the demand volatility. Professional planners will want their forecast to be as accurate as possible and while it seems to be reasonable to trust a model that has proven to be more accurate, on average, than traditional methods, high accuracy alone is not enough to realize such responsible decision making. The critical function of demand forecasting involves dealing with the many uncertainties and demand planners who are forecasting experts themselves justifiably expect to understand why the forecasting model output would be any better than their experienced intuitive instincts. Planners and managers also want to know how the machine making the forecast came to its conclusion and what are the factors that have been considered. Such insights would allow the planner to further refine and control the forecast. This is especially important when the planner feels that the machine is not systematically capturing crucial factors. So, in order to realize this, these are the concerns that we need to address:

1) Comprehensibility: Help users understand the mechanics of the model and know what is taken into account.
2) Trust: Ensure that users can trust the output of given model
3) Control: Giving users to ability to select among multiple choices based on simple explainability without requiring any familiarity with the underlying mathematics.

3 Explaining our forecast

In this paper we aim to develop and evaluate a novel method for explaining the output of a demand forecasting system through textual data. We leverage two orthogonal aspects, namely, i) temporal sentiment and, ii) topic models, to provide a better understanding of AI forecasting models.

3.1 GMI Data

In our case study we have leveraged a dataset called global market intelligence (GMI) data that contains the comments from internal sales/operations managers from all over the world, across a wide range of product families, industry sectors and business divisions. This database is updated monthly by sales and operations managers across: i) geographical regions, ii) business sales division, iii) target market (ICT, Automotive, etc.). For each entry there is a comment and a corresponding score assigned by the manager that captures the sentiment on a scale of 1 to 5.

3.2 Sentiment Analysis

Sentiment analysis is the process of using natural language processing (NLP) to analyze market and customer sentiment. It is well used in business intelligence as it enables the user to understand the subjective reasons why the market is behaving in a certain way. Leveraging such information from our GMI data can give us valuable insights when providing explainability.

3.2.1 Topic-wise Sentiment modelling

We chose to leverage the LSTM[2] algorithm for our sentiment classification task because it models the sequential dynamics of data and thereby is effective for our application. We train our LSTM model based on a custom dictionary generated from our GMI database. This is to ensure that the vocabulary includes domain knowledge and reflects the contextual embedding of the words typically used by sales managers and their domain characterization.

The workflow of the system is described in [Fig.2]. First, we collect tokens from the input corpus which is the set of all GMI comments and perform basic preprocessing (stop word removal, lemmatization, etc.) to create the vocabulary. Next, we create a dictionary based on this vocabulary and transforms each comment to a numerical feature to create the corresponding vectors. Thereafter, the examples in the training set are used to train the LSTM model and the same model is used to predict the sentiment on unseen comments.

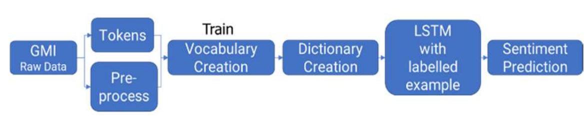

Fig. 2: Sentiment analysis model with custom dictionary

Furthermore, in order to give more consistency to our sentiment analysis, we develop a novel method that integrates the concept of topic modelling with the pure sentiment information. This is essentially done to track temporal dynamics of the sentiment topic wise. We use conventional topic modelling algorithms (BERTopic[3]) and calculate a topic probability for each GMI comment. As the demand forecasting is occurring on a monthly basis, we aggregate said topic probability and calculate a monthly score.

Let N be the number of comments in the GMI database and let K be the number of topics decided by the user. They are indexed by i and k, respectively. We define $r(i)$ as the sentiment score computed by the sentiment classification model mentioned above. Let $\alpha(i, k)$ be the probability of topic k in document i, where $\sum_{k=1}^{K} \alpha(i,k) = 1$ and $\alpha(i,k) \geq 0$ for all i, k. We define a topic mood score that reflects the macro level sentiment corresponding to a topic based on the whole corpus. Each document i is classified to k for which $\alpha(i, k)$ is maximum. Within that topic the topic mood score is defined as:

$$\beta(k) = \frac{\mu(+,k)}{\mu(+,k)+\mu(0,k)+\mu(-,k)} \qquad \text{Eq.(1)}$$

where $\mu(+, k), \mu(0, k)$ *and* $\mu(-, k)$ are number of positive, neutral, and negative comments for topic k with maximum $\alpha(i,k)$ among all k in {1, K}. Finally, the Topic wise sentiment score for each i in GMI, denoted as S(i) is computed as:

$$S(i) = r(i) \sum_{k=1}^{K} \alpha(i,k)\, \beta(k) \qquad \text{Eq.(2)}$$

This is one of the novel contributions of this paper. Now, to aggregate this monthly we took the median S(i) for all the comments from said month. This reflects the temporal dynamics of the linear combination of topic-wise sentiment and per document topic content and eventually aggregated over months.

3.2.2 Evaluation

For our experimental evaluation we use the GMI dataset that contains a total of 785 comments out of which 392 are positively rated and 393 are negative. We use 80% for training and evaluate on the remaining 20% . Our model calculates the topic-wise sentiment score of a given salesperson comment and classifies it as a positive or negative sentiment.
Results are summarized in Table1.

Table 1: Sentiment Analysis Results on GMI data

	Precision	Recall	F1Score	Support
neg	0.76	0.59	0.66	80
pos	0.66	0.81	0.72	78

3.3 Explaining the forecast

We implement the sentiment analysis model mentioned in (3.2) in our forecasts system and use this new information to strengthen its explainability. In this article we use forecast for Panasonic products in the automotive market as a test case and explain our method. Here, we leverage the feature importance calculation method based on Shapley Additive explanations (SHAP) [4] which is a fairly well-developed method for feature importance analysis.

3.3.1 Case Study: Automotive market

The case study is performed with respect to PCB relay which is a product used in the automotive market. This product is selected as it has been recently predicted that the automotive market is going through a dramatic transformation. As per the current practices we are using automotive market indexes such as world-wide production data for forecasting the sales. These features are macroscopic in nature in the sense that typically they reflect the dynamics of an entire industry sector and thereby are unlikely to change drastically over a small to moderate period. Whereas the market intelligence data internally available to the company is more relevant to specific products, business divisions and geographical locations. We aim to leverage that kind of internal dataset (mostly textual) to explain the forecast to facilitate large scale adoption among different functions within the company.

We collected the GMI data from August 2021 to May 2022 and used a 3 months' time shift to make sales forecast for the product mentioned above. We compute and evaluate the forecast generated for the month of July 2022. We produced the forecast under two different feature sets, i) conventional automotive market indexes, ii) features generated from GMI database. Lastly, we created features through topic models for 4 independent topics following the method developed in (3.2) and computed the forecast with all these above-mentioned features.

3.3.3 Evaluation

In order to evaluate the effectiveness of the framework developed we compare it with the conventional case where the forecast is generated through relevant macroeconomic indices.

The traditional features used to forecast the sales of PCB relay are mainly the features generated from the automotive global market indexes such as production figures at different regions. We use the tool called TSFresh[5] to generate the statistical characteristic from these timeseries on automotive market index production volume at different regions. We replace this set of features by features obtained from GMI by applying NLP tools described in Section (3.2). It is found that the difference in forecast generated by two different sets of features is only 3.17%. The forecast generated by traditional features and textual features are reasonably close to each other meaning former can be explained by the later with no compromise in error. However, the feature importance analysis through textual features is much more explainable and interpretable by business users with varying degrees of domain, product, and industry knowledge. [Fig.3] presents the forecast generated by our model using textual features. It is observed that a decrease in sales is being predicted for the month of July when compared to the last few months, and this can be very well explained using our proposed method when using textual features.

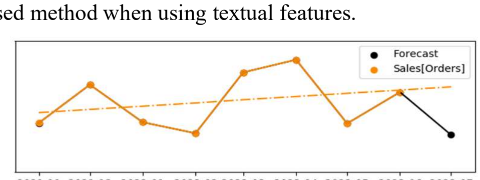

Fig. 3: Forecast Result.
A decrease in sales is predicted for July2022

First and foremost, the traditional features are macroscopic, whereas the numerical features are very complex statistical functions of microscopic factors. As a reason it is practically impossible to explain the forecast using only the traditional features. Moreover, they are often ambiguous and contrary to general wisdom. On the other hand, the textual features are more microscopic and contain information specific to the Panasonic business as well as to relevant market segments.

In our example, two complex functions of different quantiles and a reversed time asymmetry statistics of the automotive market index timeseries turns out to be three top features in the traditional forecasting method [Fig.4]. On the other hand, in our proposed method, the number of low (=2 out of 5) sentiment markets, Automobile Segment Mood and Automobile Market Sentiment turns out to be the most influencing features. The first feature is the number of comments where the calculated score equals to 2, which is intuitively a negatively correlated feature with reference to sales. 'Automobile Segment Mood' is the automotive sentiment calculated with reference to the overall market aggregated monthly. More precisely, this is S(i) in [Eq.(2)] for the comments where the topic automotive is predominant for each month and since we do this topic wise segmentation it gives additional insights. This is nonetheless very in line with the nature of the forecast target product. The third most important feature is the r(i) as described in [Eq.(2)] for the comments where is the automotive topic is predominant.

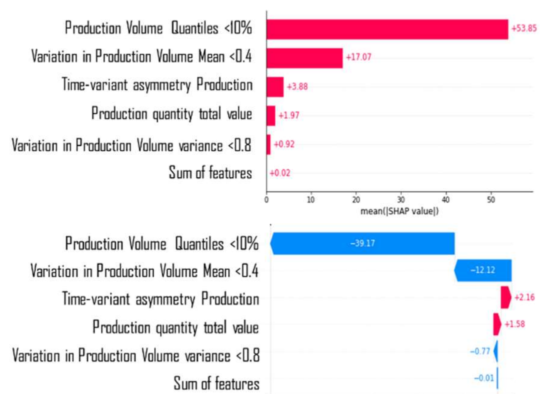

Fig. 4: Feature importance analysis of the forecast produced by traditional features

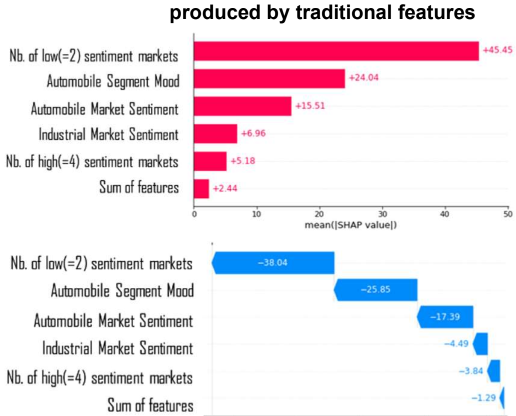

Fig. 5: Feature importance analysis of the forecast produced by textual features.

It is to be noted that we created the features for all the topics including Semiconductor, Automotive, Industrial and others out of which only the ones related to automotive resulted in higher SHAP score than others. The global feature importance analysis ([Fig.4] Upper & [Fig.5] Upper) explains the model over the entire training dataset whereas the local importance analysis ([Fig.4]lower & [Fig.5] lower) explains the forecast for a particular month. We observe that the top 5 features for globally and locally important feature are same for both traditional and textual features, respectively.

The top important feature is Number of comments with low score (2 out of 5) of which the sales should be a decreasing function. As shown in the local feature importance figure the number of low sentiment score decreases the forecast by 38 units. This is because the average value of this feature is 6 and for the month of July this is 9 which is much larger than the average value. Similarly for the second most important feature, namely, Automobile Segment Mood takes the average value of 0.12, whereas for July it takes a value of 0.08. The sales are an increasing function of this feature and thereby the lower-than-average value of this feature decreases the forecast by 26 units. The above provides a much clear and better explainability of AI forecast to the user, which is the main objective of this paper.

We have presented this concept to internal sales departments and have received highly rated feedback. Our concept was particularly praised for it can be used to estimate domain specific information that can help them plan, operate and make the right informal decisions. We were also encouraged to expand our findings to other crucial markets such as ICT which are much more complex and where insightful information is much harder to come by.

4 Conclusion

We have developed a demand forecasting model for PSI optimization, and it has been implemented into multiple sales department's monthly process. Furthermore, in this paper we introduced a novel method that provides a better understanding and explanations on why each forecast was made. We presented a method based on the popular technique (SHAP). Our high-level objective is to help users make informed decisions and avoid errors that could result in significant losses. Such explanations reinforce our AI models, increase confidence in the system, and help identify errors and performance problems. In future works we plan to explore higher level of explanation based on textual data such as the creation of causal chains and economic knowledge graphs specific to the Panasonic industry business.

References

1) Miura K. (2020). Hierarchical Forecasting AI for Global Device Demand, Panasonic Symposium2020
2) Hochreiter, S. (1997). Long short-term memory. Neural Computation, 9(8)
3) M.Grootendorst (2022). Neural topic modeling with a class-based TF-IDF procedure, arXiv:2203.05794
4) Scott Lundberg. , & Co. (2017). A Unified Approach to Interpreting Model Predictions, NIPS2017.
5) Christ, M. (2018). Time Series Feature Extraction on basis of Scalable Hypothesis tests. NC 307 (2018) 72-77