Exploring Language Models to Analyze Market Demand Sentiments from News

Abir Naskar, Manjira Sinha and **Tirthankar Dasgupta**
TCS Research
India
(abir.naskar, sinha.manjira, dasgupta.tirthankar)@tcs.com

Abstract

Obtaining demand trends for products is an essential aspect of supply chain planning. It helps in generating scenarios for simulation before actual demands start pouring in. Presently, experts obtain this number manually from different News sources. In this paper, we have presented methods that can automate the information acquisition process. We have presented a joint framework that performs information extraction and sentiment analysis to acquire demand related information from business text documents. The proposed system leverages a TwinBERT-based deep neural network model to first extract product information for which demand is associated and then identify the respective sentiment polarity. The articles are also subjected to causal analytics, that, together yield rich contextual information about reasons for rise or fall of demand of various products. The enriched information is targeted for the decision-makers, analysts and knowledge workers. We have exhaustively evaluated our proposed models with datasets curated and annotated for two different domains namely, *automobile* sector and *housing*. The proposed model outperforms the existing baseline systems.

1 Introduction

Demand forecasting is one of the fundamental aspects of business planning that drives a host of strategic and operational decisions taken by a company. It feeds into budgeting, financial planning, campaign management for sales and marketing divisions and capacity planning through scenario generation. Different stages of supply chain planning leverage information from different sources to predict probable demand for products or services across different regions. These include current sales data as well as insights about future demand gathered from a plethora of sources like social media, customer surveys or analyst reports, News etc. While analysis of relevant consumer generated content from social media has been found to have direct impact on demand of consumer goods in short term, for long term demand assessment human experts still rely on business News and analyst reports. Given the high volumes of business relevant content available for real-time demand assessment today, human curation and compilation of such information is gradually becoming impossible. Consequently text mining methods are envisaged to play a significant role in extracting and compiling the relevant information from a multitude of sources in an efficient and effective way.

In this paper, we have presented an information extraction model that exploits the transformer based neural network architectures to acquire demand related information from business text documents like News articles and reports. The task is to first identify demand related information from News and other business text sources and then resolve all associated aspects of it like product or service names, region, time, and rise or fall of demand along with reasons of the specific rise and fall, if mentioned in the text. Text elements indicative of positive sentiments in association to demand are indicative of rise in demand while negative sentiments associated with a demand indicates a fall in demand. Since business text can be written in very complex ways, hence extracting all these parameters correctly from text is a non-trivial task. For example, Table 1 example (1) contains interesting insight about positive and negative demand information for two different car models. In this paper, we present results related to the automobile and housing sector, and show how demand related insights for ten years are generated. Performance analysis of the proposed architecture is obtained using a gold-standard data set that has been manually annotated for computing the accuracy of the information extraction and enrichment processes.

News documents and analyst reports not only

Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis, pages 264–272
August 15, 2024

Text segments with causality marked	Subject-of-Demand: Polarity(['SDE','P']
While demand for smaller model is soaring, sales of Some traditional vehicles have remained strong.	['smaller models', 'NEG'] ['traditional vehicles','POS']
The report found that out of A3 compact Q3 and Q5 SUVS of Audi, the demand for Q5 SUVs are significantly higher.	['AUDI', 'POS'] ['Q5 SUV','POS']
{Earnings of Hyundai Kia will likely drop}$_{effect}$ for the third quarter of the year *{due to slump in local production}*$_{cause}$.	['Hyundai Kia', 'NEG']

Table 1: Examples of complex text segment with demand - the target is to extract all relevant information components

carry information related to future demands, they also carry expert insights on reasons for demand fluctuations. These articles can be mined for creating a knowledge base of demand impacting factors that can be used in a predictive solution for improved demand forecasting. A wide range of events has been found to impact demand in direct and indirect ways. Economic recession, pandemic, political unrest, legal battles are just few among a large class of events that have had provable impact on demands for a wide range of products and services in recent times.

We would like to emphasize that the framework does not implement any demand-forecasting model. Rather, the proposed framework gathers early demand signals and post-facto knowledge about demand alterations and presents these insights to a human decision maker, who can further use these to refine the outputs of a mathematical forecasting model. In this work we restrict ourselves to mining demand relation information from News and analysts reports only and do not consider social media inputs. But this being a reasonably well-explored area, we do provide a review of work done in this area. We have worked with large collections of News articles to extract demand related insights.

The rest of the paper is organised as follows: section 2 discusses about the problem definition. Section 3 presents our proposed TwinBERT architecture for subject of demand identification and polarity classification task. Section 4 presents the dataset used for training and testing the models followed by the experiment design and results in section 5. Section 6 discusses about the reason behind the rise and fall of demand and finally section 7 concludes the paper.

2 Mining demand related insights from business text

Table 1 shows few example texts that contains information about demand for automobiles across the globe. As we can see, it may contain demand related information at very granular level that includes demand estimates for specific regions along with time when those demand patterns are likely to be seen, or at very high level like "passenger vehicles". For example, the text segment, *"The report found that out of A3 compact Q3 and Q5 SUVS of Audi, the demand for Q5 SUVs are significantly higher."*, mentions about three product names namely, *A3*, *compact Q3* and *Q5 SUVS* of Audi. Out of this only demand of *Q5 SUV* has been mentioned in the text with a *positive* demand and the rest are *neutral*. Therefore it is important to correctly filter out the exact subject of demand entity along with its sentiment for which demand is associated. We define the Subject-of-Demand Entity (SDE) as: *An entity with whom information components exhibit a direct relation by virtue of being linked semantically*. For Example, in the text, "*With increased demand Mahindra is forced to enhance production capacity of Quantoas Nasik plant.*", "*Quanto*" is the SDE.

Apart from extracting the subject-of-demand and their corresponding polarities, it is also important to gather key insights about the reason behind the rise or fall of demand. Such insights can be identified from text documents by performing proper causal analysis of a given text input. For example, Table 1 example 3 illustrate a negative demand for *Hyundai KIA* with proper causal reasoning (*"slump in local production"*).

Based on the above factors, we define the following target tasks for this paper:

- Task-1: Given a text mentioning about a product demand, identify the relevant Subject-of-Demand entities (SDEs).
- Task-2: Given a text with identified SDEs, identify the sentiment polarities (*positive* for rise in demand and *negative* for fall in demand) corresponding to each SDE.

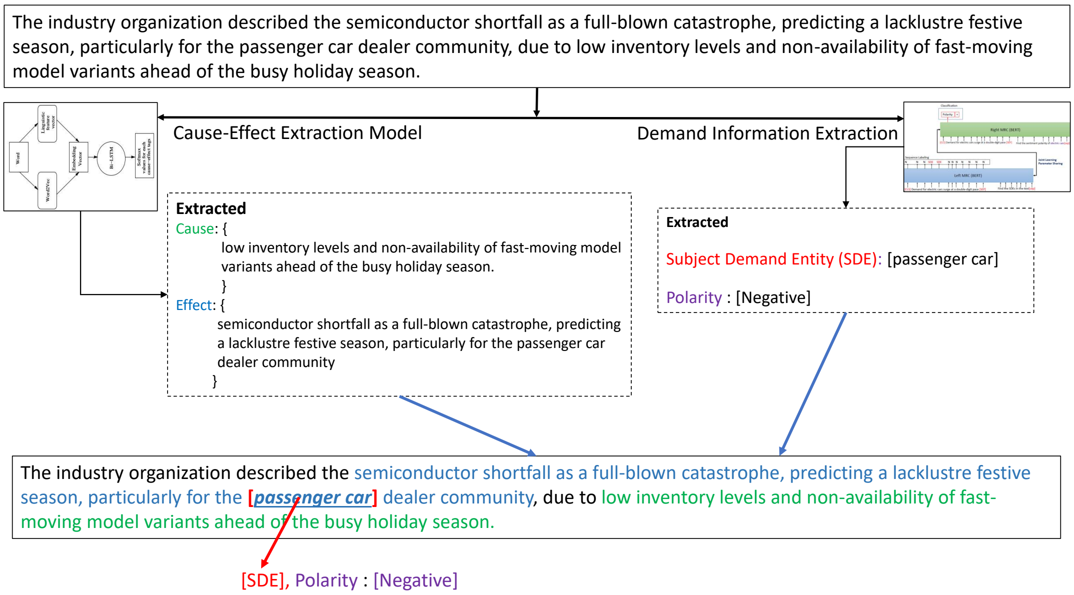

Figure 1: Working of the demand and causality extraction engine. The same text is passed to both the demand extraction as well as causality identification unit. The output label of both the models are merged together to get the final output.

- Task-3: Given a text with demand information, identify the specific causes of rise and fall of demand.

Accordingly, in this paper, we propose a joint training framework to extract both SDE as well as the sentiment corresponding to the extracted SDE from textual mentions. We use BERT (Devlin et al., 2018) as our backbone network and use a sequence labeling model to detect the start/end positions of SDE-sentiment pair(s) from a text segment.

Following the work of (Mao et al., 2021), we propose a TwinBERT architecture to represent the above extraction task using two machine reading comprehension (MRC) problems. MRC methods are known to be effective if a pre-trained BERT model is used. We decompose the SDE-polarity pair extraction task to two different sub-tasks of SDE detection ($BERT_1$)and sentiment identification ($BERT_2$).

Similarly, for causal inference, we have used the cause-effect extraction tool as presented in the literature (Dasgupta et al., 2018). The proposed model is based on linguistically informed BiLSTM architecture (LiBiLSTM) to extract cause and effect events from a given input text.

As illustrated in Figure 1, the same input text segment is passed to both the TwinBERT architecture as well as the cause-effect extraction architecture. The output of both the models are then combined together to get the respective *SDE, Polarity and causes-effect* relations. In the next section we will define in details the working of the TwinBERT model for SDE extraction and polarity classification task.

3 The proposed TwinBERT model for subject-of-demand extraction and polarity classification

As illustrated in Figure 2, our model consists of two parts. Both parts use a multi-layer bidirectional Transformer based language representation model (BERT) (Devlin et al., 2018) as their backbone models to encode the context information. The goal of the left part is to extract all SDEs from the given text. As we discussed earlier, we have used the sequence labeling task for this purpose. The goal of the right part is to extract the sentiment polarity with respect to a given specific SDE. This is done by applying a classification model that classifies a given text segment based on its [CLS] token. It is worth mentioning here that an input text may have multiple SDEs. Therefore, the right part of the model will take input separate text-SDE pairs corresponding to each SDE extracted by the left part.

To obtain the sequence classification and sequence labeling, the final layers of the the proposed left and right TwinBERT models have been

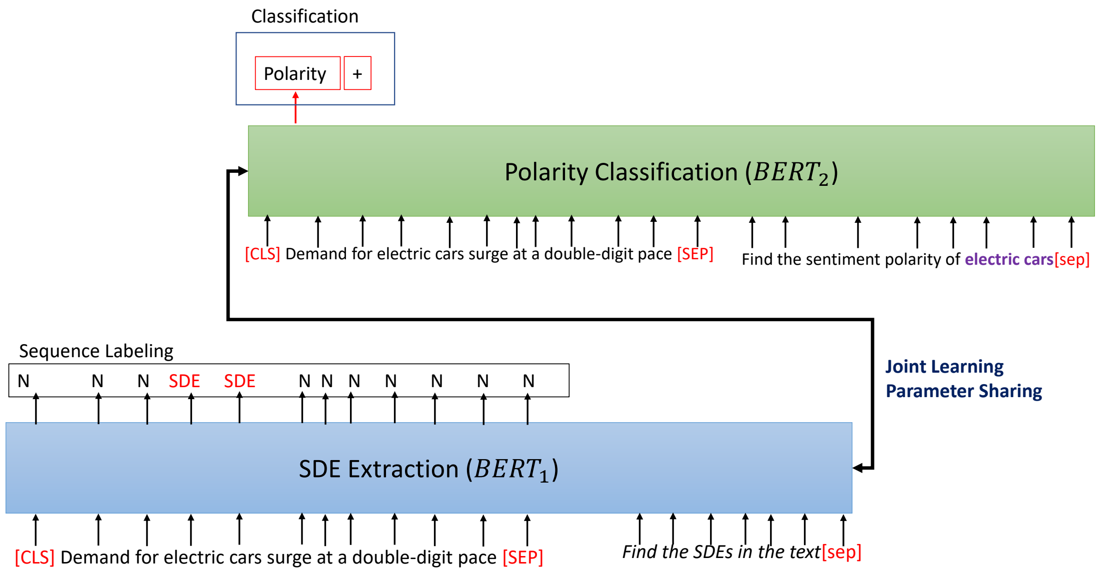

Figure 2: Overview of the Joint Training of Subject-of-Demand (SDE) Extraction and SDE Polarity Classification.

trained with two separate loss functions L_1 and L_2. Where, $L_1(\theta) = -\sum_{t=1}^{M}\sum_{k=1}^{K}\bar{y}_t^k log(y_t)$ and $L_2(\theta) = -\sum_{t=1}^{N}\sum_{j=1}^{J}\bar{q}_t^{i,j} log(q_t^i)$ q_t is the vector representation of the predicted output of the model for the input word w_t^i. K and J are the number of class labels for each task (which is three in our case). The model is fine-tuned end-to-end via minimizing the cross-entropy loss.

We define the joint loss function using a linear combination of the loss functions of the two tasks as:

$$L_{joint}(\theta) = \lambda * L_1(\theta) + (1-\lambda) * I_{[y_{text}==1]} * L_2(\theta) \quad (1)$$

Where, λ controls the contribution of losses of the individual tasks in the overall joint loss. $I_{[y_{text}==1]}$ is an indicator function which activates the loss only when the corresponding *SDE-Extractor* classification label is 1, since we do not want to backpropagate the *PolarityClassifier* loss when the corresponding *SDE-Extractor* output is 0 i.e if the *SDE-Extractor* does not return any SDE as output.

4 The dataset

We have curated around 74,150 news documents across the two target domains over the time period of ten years(2012 to 2021). From the given dataset we have extracted all the text segment containing words related to the concept "demand", using seed words like: *demand, requirement, need, market need, desire* etc. This gave us around 29000 text documents that contain demand related concepts. For example, text segment (1) below is mentioning about demand of automobile products where as text (2) is not related to any product demand.

1. *November saw demand for new cars, trucks and crossovers surge at a double-digit pace.*

2. *There have been numerous reports of police officers stopping such cars, and demanding that the driver produces his or her tax registration ID.*

Once, we filter out the product demand text, we present them for expert level annotations across each domains by six annotators. The annotation process undergoes the following tasks:

Task-1: Given a demand related textual mention, identify the respective subject-of-demand entities (SDE).

Task-2: If the text contains at least one SDE, then the task is to determine the sentiment polarity associated with the entity. The task here is to classify the sentiment as positive, negative or neutral. Some sample annotations are mentioned in Table 1. Finally, we have a gold standard data of 12200 text documents. Out of this, 7400 texts are from Automobile domain and the rest 4800 from Housing domain. Overall, around 43% positive demand samples, 37% negative demand samples and 20% neutral demand samples. We have used 70% of the overall data for training the classifier and rest 30% for testing purpose.

4.1 Conversion to *TwinBERT* compatible data format

As illustrated above, the original annotated dataset needs to be converted before it is fed into the *TwinBERT* Network. Both the *SDE-Extractor* and the *PolarityClassifier* use the input text as their contexts along with a specific query. The *SDE-Extractor* is constructed with the query, q1 = "Find the SDE terms in the text." While, the *PolarityClassifier* is constructed with the query q2 = "Find the sentiment polarity for the <SDE> in the text." This is illustrated in Figure 3.

5 Evaluation

In order to demonstrate the importance of the proposed neural network architecture for demand extraction, we make a comparative study of the performances of the model with respect to other standard neural network architecture. We have kept the same set of hyper-parameters for understanding the difference in their performance. The experiments were conducted using the following models: 1) Cascaded CNN-BiLSTM model: We use a standard CNN model coupled with a Bilstm layer for the extraction of SDEs and further classification of the aspect level polarities. We used the pre-trained Word2Vec embeddings to train the model. We run for training using a mini-batch size of 128 for each fold, and optimized using the Adam Optimizer. For the Bi-LSTM, 64 hidden units were used. For the CNN, layers for kernel sizes 2 to 6 were included in the network. 2) BERT model: We used the pre-trained BERT model as proposed by (Devlin et al., 2018) and fined tuned them over the proposed dataset. Here, two separate models are used to train the SDEs and their polarities. In all our experiments, 10-fold cross validation was used for the purpose of fair evaluation on the datasets. For each fold, 10 epochs were run for training using a mini-batch size of 12 for each fold, and optimized using the Adam Optimizer with learning rate of $2 * e^{-5}$.

Evaluating the Subject-of-Demand Entity Extraction: We quantify the performance of the demand classification score in terms of the precision, recall and F-measure values. For the SDE extraction task we perform the evaluation with respect to the different neural network models as discussed in the previous section. The extracted entities were then compared with the gold standard annotations. Table 2 depicts the evaluation results of the SDE extraction and polarity classification model for both Automobile and Housing domain data. For *Automobile domain*, we have achieved an F-measure of 0.81 with precision of 0.79 and recall of 0.83 respectively for Subject-of-Demand entities. Similarly, for the *Housing domain*, we have observed an F-Score of 0.79 with a precision of 0.77 and recall of 0.81.

Evaluating the sentiment polarity of extracted SDEs classifier: We quantify the performance of the polarity classification score in terms of the precision, recall and F-measure values. Table 2 depicts the evaluation results of the demand classification system for both *Automobile* and *Housing* domain data. We found that throughout all the target classes the performance of the TwinBERT network is significantly higher than the individual single task BERT as well as multi-task BERT based models. The proposed architecture significantly reduces the false negative score and achieves a high true positive score, thereby achieving a high precision and recall. During the analysis of the individual datasets we have observed that for the Automobile dataset, we have achieved an F1 score of 82% using the TwinBERT model. This is the highest accuracy that we have achieved between both the datasets. In around 20% of cases our system failed to classify the demand class correctly. For Housing domain, we achieved an accuracy of 71.2% with precision of 73%, recall of 79% and F-measure of 75.8%.

5.1 Comparison of Proposed Model Architecture With LLMs

In the era of large language models (LLMs), there is considerable potential to outperform numerous transformer designs. After conducting the experiments on the given dataset, we compared its output to that of LLAMA-2 13B (Touvron et al., 2023) and Mistral-7B (Jiang et al., 2023). First, we have evaluated the SDE detection ability of LLAMA-2 using zero-shot (Wang et al., 2019) and few-shot prompt techniques. Here, we have used the few-shot technique demonstrated by (Min et al., 2022) and given examples of sentences with and without SDEs as prompt.

Similarly, we have also fine-tuned the pre-trained Mistral-7B Model with the given dataset to compare LLM's ability to perform the domain-specific task of SDE detection and polarity classification with our proposed architecture. The Mistral-7B outperforms the LLAMA-2 34B despite having

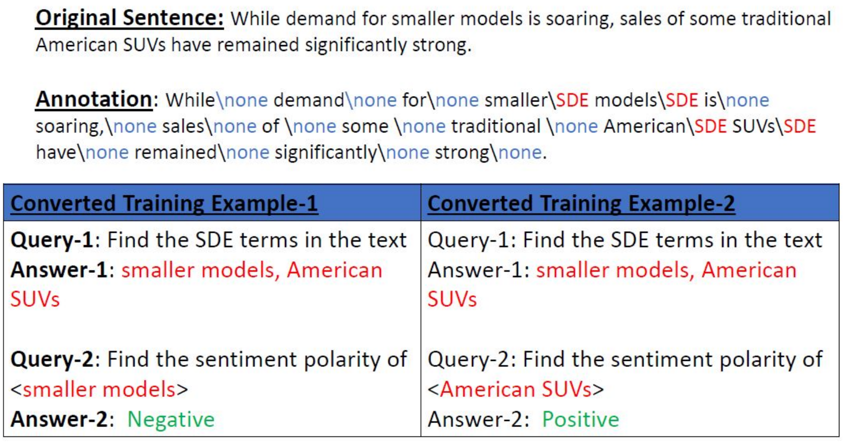

Figure 3: Illustration of the TwinBERT compatible data conversation

Automobile Domain												
	SDE Extraction			Positive			Negative			Neutral		
Models	**P**	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**	**P**	**R**	**F1**
$BERT_{base}$	0.62	0.66	0.64	0.59	0.77	0.67	0.63	0.77	0.69	0.64	0.71	0.67
LLAMA-2-13B	0.68	0.76	0.72	0.69	0.67	0.68	0.63	0.71	0.65	0.76	0.71	0.74
Mistral-7B	0.79	0.77	0.78	0.77	0.72	0.75	0.72	0.78	0.75	0.76	0.78	0.77
Proposed TwinBERT	0.79	0.83	0.81	0.72	0.70	0.71	0.63	0.69	0.66	0.84	0.82	0.83
Housing Domain												
$BERT_{base}$	0.69	0.66	0.67	0.62	0.71	0.66	0.52	0.68	0.58	0.74	0.67	0.70
LLAMA-2-13B	0.57	0.72	0.64	0.52	0.68	0.59	0.66	0.68	0.67			
Mistral-7B	0.73	0.78	0.75	0.64	0.74	0.69	0.59	0.71	0.650	0.67	0.69	0.67
Proposed TwinBERT	0.77	0.81	0.79	0.67	0.77	0.72	0.62	0.68	0.65	0.64	0.69	0.66

Table 2: SDE extraction and polarity classification results for the automobile and housing domain.

only 7.3 billion parameters on various benchmarks (Jiang et al., 2023). Here, we have primarily used transfer learning, with additional modifications such as quantization and the integration of LoRA adapters (Dettmers et al., 2023) to fine-tune Mistral. The training process involves several key steps. Each data sample is augmented with a prompt indicating the task context and the statement to be evaluated for SDE extraction and polarity classification.

5.2 Outcome of fine-tuned Mistral-7B and LLAMA-2

Empirical evidence presented in Table-2 demonstrates that our TwinBERT architecture achieves superior performance compared to LLAMA-2 13B. As we can see, the performance of LLAMA-2 using the few-shot approach was notably limited. This limitation stemmed from the complexity of defining SDEs, which necessitates a comprehensive representation beyond the provided four examples as prompt. As evidenced in the presented table (Ref: Table 2), while LLAMA-2 achieved a high precision score, its recall and F1 scores were significantly lower, primarily due to its tendency to classify the majority of sentences as not-claims. Consequently, LLAMA-2 exhibited suboptimal classification performance, particularly in the zero-shot scenario. Conversely, although TwinBERT emerged as a superior classifier in Precision, Recall, and F1 scores, its superiority can be attributed to its adherence to the intricate definition of SDEs, thereby underscoring its effectiveness in classification tasks.

Here, we ran an experiment to compare the output of our proposed architecture with our fine-tuned Mistral Model. We gave the trained Mistral model 50% of the total data sample and asked it to extract the SDEs and its polarity sentences. However, the Large Language Model's hallucinatory property posed a challenge. Out of the 50% sentences, the trained Mistral Model provided a distinct classification for only 25% sentences, while the remaining 25% cases resulted in a rather confusing answer. Among those, it categorized correctly for 22% cases. Therefore, we concluded that while training the large language model on a specific domain can improve its SDE extraction capacity, the inherent property of the Large Language Model can still pose a challenge.

6 Analyzing the rise and fall of demand

Apart from classification and extraction of demand related events and entities it is useful to visualize these demand information and perform causal analysis to generate reports on the various product trends. We illustrate this considering examples from the automobile sector. Accordingly, we have crawled around 5000 automobile News articles from India and United States during the period of 2012 and 2020. Average document size ranges to 300 $\pm$ 80 words. We have applied our TwinBERT model to extract demand information such as product names, location and polarity (i.e positive and negative demand). Once the information is extracted, the different automobile model names are segregated and grouped together according to the categories. The mapping between the model names to the respective categories are done using the state-of-the-art automobile ontology auto.schema.org[1]. The automotive extension of schema.org [2] stores the most important real-world objects related to popular vehicles like cars, buses and two wheeler vehicles. While the extension allows for a fair description of all kinds of vehicles, it focuses predominantly on passenger automobiles from the retail market perspective.

Next, we pass the text segments of each articles to a causal analytics module that identifies causal events from text. For this, we have used the cause-effect extraction tool as presented in the literature (Dasgupta et al., 2018). The proposed model is based on linguistically informed BiLSTM architecture (LiBiLSTM) to extract cause and effect events from a given input text. For example, in the following text:

"In the passenger vehicle segment, showroom sales declined by 11% year on year to 243183 units as customers stayed away from the showrooms due to lack of improvement in availability of credit or finance options, higher cost of ownership and overall slowdown in the economy." The extracted cause and effect events are:

`CAUSE`: *lack of improvement in availability of credit or finance options, higher cost of ownership and overall slowdown in the economy.*

`EFFECT`: *showroom sales declined by 11%*

Figure 4 depicts the overall demand distribution of top 5 automobile types in India. During the period of 2012 to 2018 we observe a steady demand for almost all segment of vehicles. As expected, the highest demand are for two-wheeler vehicles as compared to all other types. We primarily observe two major spikes in the dataset. a) With respect to passenger vehicles we observe sudden spike in demand during 2015 particularly for fuel efficient cars and electric vehicles. b) A severe decline of demand is observed across all automobile segments during late 2018 till 2020. We try to perform an in-depth analysis of the reasons for such a rise and fall of demand. Our causal analysis during those period revealed that the sharp rises in demand during 2015 coincided with news reports of launch of hybrid, electric and fuel efficient cars of Nissan and Ford during that time period. On the other hand, major reasons for fall in demand can be attributed due to many factors including: *Overall price rise across automobile sector, Economic slowdown in India, High oil price, BS-IV implementation, and Unavailability of semiconductors.*

In Figure 5, we present the distribution of demand of the top five car models during 2012 to 2016 across the region of the Unites States of America and Canada. We observe that, while the demand for fuel-efficient vehicles increased initially, the demand trend of such cars have consistently been downwards after the initial two years. One of the primary reason behind such an event is the lowering of gasoline prices and increase in oil prices in those areas. Unsurprisingly, demand for diesel vehicles have remained consistently low in the entire region of the USA and Canada. We also observe a constant low demand for electric cars as compared to standard SUVs. Corresponding news documents indicate that the primary reason behind this is the high cost of rechargeable batteries, low gas price and limited mileage offered by electric cars. However, the demand for electric vehicles have consistently risen over the past seven years across geographic locations. Another notable observation is the exponential increase in demand for SUV cars and trucks throughout the given time period.

7 Related Works

Demand forecasting is an important aspects in the supply chain business. The demand forecasting models traditionally uses features like seasonality of goods, price points, previous experience etc. Beyer et al. (Beyer et al., 2005) explored

[1] https://schema.org/docs/automotive.html
[2] www.auto.schma.org

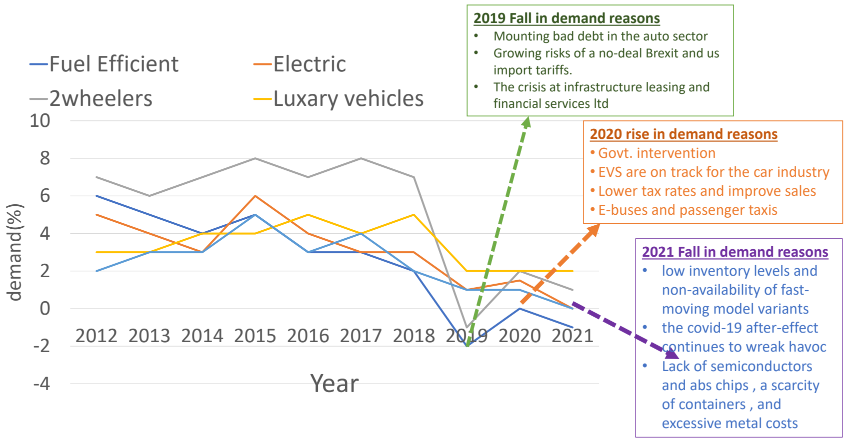

Figure 4: Demand trend analysis along with causal reasoning of automobiles in India.

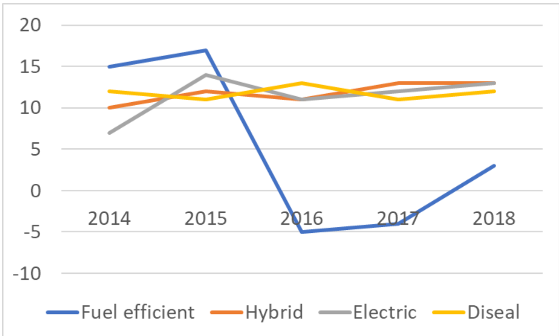

Figure 5: Demand trend analysis of top 4 car types over the past five years at USA.

the way to find demand profile of a new product which would be introduced in market using the demand profiles of similar products already in market. Berry et al. (Berry et al., 2004) examined the situation when the number of observations is associated with the number of products within a given market. It is necessary for manufacturers to give the retailers view about the demand potential of their new products. Desai et al. studied how a high-demand manufacturer can use advertising, slotting allowances, and wholesale prices to signal its high demand to retailers (Desai, 2000). Mark E. Ferguson explored statistical methods for estimating demand with constrained data and product substitutions (Ferguson, 2020). Abbasimehr et al. proposed multi-layer LSTM networks for predict the demand (Abbasimehr et al., 2020). Gunter et al. (Gunter et al., 2020) explored the Airbnb demand to New York City by employing spatial panel data at the listing level. Along with the traditional features two new features Item categorization using word2vec with clustering and session of the day based on the time was proposed by Dholakia et al. (Dholakia et al., 2020) to obtain an improved and intuitive demand forecasting model.

8 Conclusion

In this paper we have developed methods to mine product-specific demand information components from large volumes of text data. The work primarily focuses on analyzing text documents and extracts specific Subject-of-Demand Entities from text segments that mentions about demand of a product, determine the sentiment, in terms of rise and fall of demand, associated to the subject-of-demand and finally analyzes the reason for the rise and fall by performing causal analytics. Accordingly, we propose the use of a TwinBERT architecture for the entity extraction and sentiment classification task. We have evaluated our system using a manually annotated gold standard dataset belonging to two different domains. We have observed that our proposed model significantly outperforms the existing baseline models.

Limitations

1. One of the limitation of the present work is that it fails to identify implicit demands.

2. An in-depth analysis of the performance of LLMs on demand mining is required.

3. Evaluation of the causal models are not explicitly discussed in the paper.

References

Hossein Abbasimehr, Mostafa Shabani, and Mohsen Yousefi. 2020. An optimized model using lstm network for demand forecasting. *Computers & Industrial Engineering*, page 106435.

Steve Berry, Oliver B Linton, and Ariel Pakes. 2004. Limit theorems for estimating the parameters of differentiated product demand systems. *The Review of Economic Studies*, 71(3):613–654.

Dirk M Beyer, Fereydoon Safai, and Farid AitSalia. 2005. Profile-based product demand forecasting. US Patent 6,978,249.

Tirthankar Dasgupta, Rupsa Saha, Lipika Dey, and Abir Naskar. 2018. Automatic extraction of causal relations from text using linguistically informed deep neural networks. In *Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue*, pages 306–316, Melbourne, Australia. Association for Computational Linguistics.

Preyas S Desai. 2000. Multiple messages to retain retailers: Signaling new product demand. *Marketing Science*, 19(4):381–389.

Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. 2023. Qlora: Efficient finetuning of quantized llms. In *Advances in Neural Information Processing Systems*, volume 36, pages 10088–10115. Curran Associates, Inc.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Rishit Dholakia, Richa Randeria, Riya Dholakia, Hunsii Ashar, and Dipti Rana. 2020. Cognitive demand forecasting with novel features using word2vec and session of the day. In *Modern Approaches in Machine Learning and Cognitive Science: A Walkthrough*, pages 59–72. Springer.

Mark E Ferguson. 2020. Estimating demand with constrained data and product substitutions. In *Channel Strategies and Marketing Mix in a Connected World*, pages 1–27. Springer.

Ulrich Gunter, Irem Önder, and Bozana Zekan. 2020. Modeling airbnb demand to new york city while employing spatial panel data at the listing level. *Tourism Management*, 77:104000.

Albert Q. Jiang, Alexandre Sablayrolles, Arthur Mensch, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Florian Bressand, Gianna Lengyel, Guillaume Lample, Lucile Saulnier, Lélio Renard Lavaud, Marie-Anne Lachaux, Pierre Stock, Teven Le Scao, Thibaut Lavril, Thomas Wang, Timothée Lacroix, and William El Sayed. 2023. Mistral 7b. *Preprint*, arXiv:2310.06825.

Yue Mao, Yi Shen, Chao Yu, and Longjun Cai. 2021. A joint training dual-mrc framework for aspect based sentiment analysis. *arXiv preprint arXiv:2101.00816*.

Sewon Min, Xinxi Lyu, Ari Holtzman, Mikel Artetxe, Mike Lewis, Hannaneh Hajishirzi, and Luke Zettlemoyer. 2022. Rethinking the role of demonstrations: What makes in-context learning work? *Preprint*, arXiv:2202.12837.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. 2023. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*.

Wei Wang, Vincent W Zheng, Han Yu, and Chunyan Miao. 2019. A survey of zero-shot learning: Settings, methods, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 10(2):1–37.