Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 140 (2018) 170–178

Procedia Computer Science

www.elsevier.com/locate/procedia

Complex Adaptive Systems Conference with Theme: Cyber Physical Systems and Deep Learning, CAS 2018,
5 November – 7 November 2018, Chicago, Illinois, USA

Customer Perception Analysis Using Deep Learning and NLP

Sridhar Ramaswamy*, Natalie DeClerck

**Analytics Steward, Caterpillar Inc., 325 S Route 31, Montgomery, IL 60538*
Data Scientist, Caterpillar Inc., 501 SW Jefferson Ave, Peoria, IL 61605

Abstract

Understanding customer behavior and driving customer satisfaction are necessary for any business to succeed in the competing market. Companies need to be aware of the prevailing customer perceptions to make more accurate and effective plans for product development and marketing. Customer feedback is available through multiple channels. Specifically, we are interested in the unstructured data available as text that would be available through social media, comments from a survey, voice recordings of customer interactions, and chat transcripts. Analyzing such data correctly is critical, as it reveals everything from buying trends to product flaws and provides a significant business advantage. It would further strengthen business opportunity to uncover customer interests, product improvements, and marketing insights. In this paper, we explore different technologies of Deep Learning and Natural Language Processing (NLP) that would help analyze better the contextual information to capture customer feedback.

© 2018 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)
Selection and peer-review under responsibility of the Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems.

Keywords: Customer Perception; Natural Language Processing; Semantic tagging; Deep Learning;

1. Introduction

In the recent years, there has been a huge impact on businesses primarily driven by, how customers perceive about their products and services [1]. Especially, in the areas of marketing, product quality and the end-to-end delivery, business is keen on understanding and reacting to customer views instantaneously. Customers convey their views

* Corresponding author. Tel.: +1 847 508 9786
E-mail address: Ramaswamy_sridhar@cat.com, sridharr24@gmail.com

1877-0509 © 2018 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)
Selection and peer-review under responsibility of the Complex Adaptive Systems Conference with Theme: Engineering Cyber Physical Systems.
10.1016/j.procs.2018.10.326

through multiple modes or channels, ranging from responses to online text messages and phone surveys, emails and social media. Their demands are much heard and have direct influence on a company's market share and longevity. Hence, it becomes inevitable to stay up with competition due to faster trends in customer behavior and loyalty. This has forced every company to closely follow the views of its customers, their concerns, likes and dislikes, and provide niche changes to their product design, services and marketing strategy. However, as a preliminary step in this evolution is the challenge of scrutinizing volumes of data; in some cases, incomprehensible unstructured texts, that portray the reasons for a possible customer churn that may otherwise go unnoticed. Furthermore, grievances, appreciation, or reasons for complaint are driven by subjectivity of expectation or satisfaction influenced by factors, that could be captured as part of the data. From business point of view, the aim is to consider all this information to understand the context and take necessary follow up actions for improving customer experience.

Towards this goal, this paper intends to explore the analytical aspects such as Natural Language Processing and Deep Learning that will derive meaningful insights from Customer Perception (CP). In this paper, we analyze the specific question: Is it possible to solely understand and capture customer concerns and reasons, besides sentiment using DL/NLP?

To explain the concept discussed here, we consider the cars reviews data from [2]. This dataset (henceforth referenced in this paper as "car reviews") contains full reviews of cars for model-years 2007, 2008, and 2009 for specific models. There are about 140-250 cars for each model year. We have considered for our analysis the text field "Favorites" which contains customer's experience of different car types and features. The aim is to understand what customer favorite features and grievances are, from general design cars perspective for improving car future. For this paper, we have not separated the dataset by car models or brand, as those will be considered for a future study.

2. Customer Perception (CP)

In the areas of brand, marketing and quality, customer views are gathered and studied for improving the business. In general, there is a lot literature of analyzing customer loyalty and customer satisfaction using statistical measures derived out of quantitative data [3,4,5,6]. The survey questionnaire is mapped to quantitative data and studied extensively. Recently, the customer views are getting gathered through many channels, some of them, not quantitative, instantaneous and in large volumes. In addition, businesses run routine surveys blind or targeted, to understand the customer experience. Furthermore, customer reviews are gathered through web, email or mobile, or claims / satisfaction questionnaires. Besides the quantitative information, the customer experiences received in terms of texts can be classified into three categories:

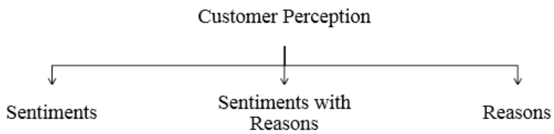


Fig. 1. Customer Perception Categories

Customer sentiments themselves are of not much help in service offering improvements or product changes. Sentiments with reasons can give insight into actionable items by an enterprise. CP reasons in the text could arise in as customer comments on product purchase, product failures or claims.

The benefit of understanding CP is, businesses can track customer loyalty or churn [7], and act upon without disruption. With the surge of social media through mobile and portable devices, the trend only will only grow for customers to provide instantaneous feedbacks. This warrants companies to constantly listen, understand and adapt to customer expectations. It has been shown that timely recovery methods can prevent customers from switching a brand or product [8]. The predicament lies in interpreting a text with no other informational cues, such as customer verbal tone or facial expressions, to make domain-relevant inferences [8]. However, in the case of surveys, unlike the customer tweets or mobile app questionnaires, survey instruments can provide additional leverage of having industry-specific questions to a comment or claim, which will provide additional insight on customer view expressed.

As for analyzing CP is concerned, there has been a vast literature of opinion mining and NLP techniques applied to customer sentiments [9,10,11]. These sentiments are useful for marketing and sales purposes. However, it would be difficult to interpret and make amendments to business, if there are no reasons attached to them. Most of the actionable interests of customer perception are derived from reasons and reviews. Here, we would like to concentrate on the text of reasons and reviews, to analyze CP. Towards this, we would like to explore NLP and deep learning concepts on how they would help in deriving insights.

3. Natural Language Modeling for CP

At first, we review the Natural Language Understanding (NLU) aspect of the NLP, for analyzing CP. In general [12], traditional NLP/NLU processing can be described as follows:

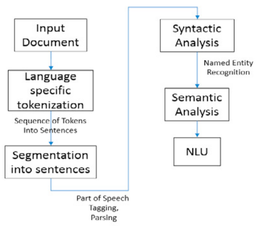


Fig. 2. Traditional NLP Pipeline

These steps are reasonable for text processing and perform well in NLP tasks of general texts. But they must be supported by business specific annotations, ontology, and semantic reasoning to make meaningful derivation. Without that, regular NLP pipeline of tokenizing, parsing, entity recognition, etc. would work only on word computations and word linkages found in documents that will lead to inconsistent retrieval of information. For example, from the car review dataset [2], let us consider a document under "favorites" reported by the customer:

Plenty of passing power and very comfortable on long trips. We need the AWD for the snow and the kids stay entertained with the AV system. Acuras are rare in Germany and people wondering what kind of vehicle I have. If you are in the market for a luxury SUV for family touring, with cool tech toys to play with, MDX can't be beat.

Fig. 3. Sample document from car reviews

In general, phrases, like "in the market" and "passing power" would be taken literally, without semantically relating to car product features and performance. Hence interpreting using basic NLP parsing or word2vec [13,14] embeddings, it is a challenge to perform an effective business analysis interpreting millions of such reviews without an understanding of key references (highlighted in blue in Figure 3).

3.1. Rule-Based Semantics Annotation

To extract syntactically actionable insights from customer feedbacks, it is necessary to understand the business intention for collecting them. Businesses run customer surveys or reviews for improvement or to proactively track customer churn, from customer complaints for products and service improvements; and in some cases, to capture the most desirable features from a product purchase or new product introduction. With better business understanding, it will be easier to create rule-based semantics relating to the problem in question [14,15].

If the rule-based semantics need to be constructed as unsupervised learning, then business documents like manuals must be scraped, or concept associations can be extracted from public sites like Wikipedia. Otherwise, Named Entity Recognition (NER) must be done on input documents [16], to get the initial understanding of concepts in them. Then using the NER, the explicit semantic representations can be used to create the Bag of Concepts (BoC) [15]. BoC refers to understanding the concept of a Bag-of-Words to its true "unit of meaning". To illustrate this using the car reviews dataset [2], we mark below the possible BoC from two documents:

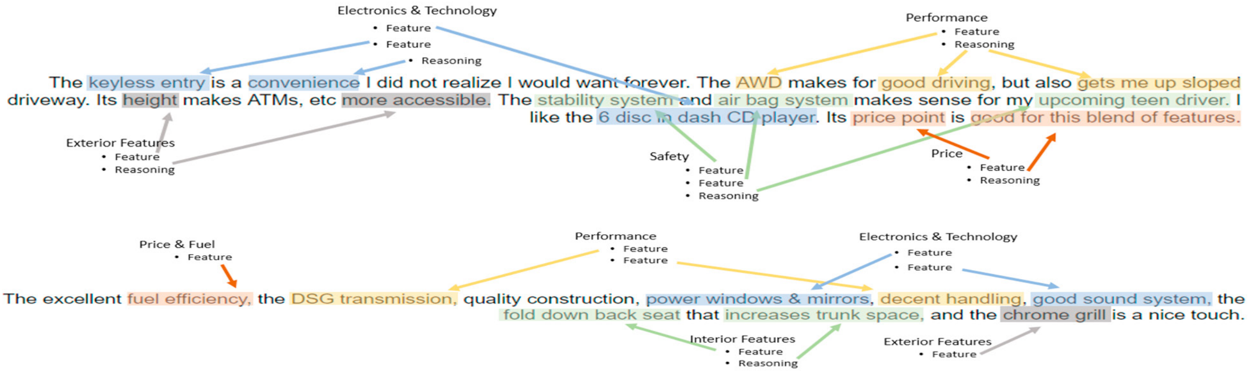


Fig. 4. Semantic tagging of documents from car reviews

The anchor texts are annotated based on the business problem of car reviews analysis. Beyond rule-based annotation, word-sense disambiguation should be sorted out manually, or by the training sequences created out of relevant document scrapings [17] using Milne and Witten algorithm [18] from Wikipedia or intranet site documents by the enterprise. The algorithm uses machine learning, candidate selection, disambiguation and link-detection. Hence, we emphasize for CP, rule-based semantic annotation plays an important role in deciphering the context of customer views. For instance, a sample semantic annotation of CP from cars review is given below:

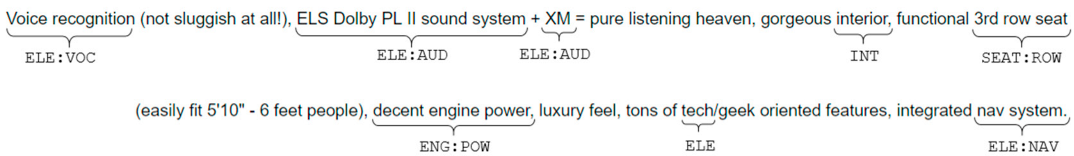


Fig. 5. Annotated sample document from car reviews

Development of business rule-based semantic annotation and product-related taxonomy [14] are essential before proceeding to other aspects of NLP. This discipline of language annotation specific to the business-driven context, serves as a critical link in developing intelligent human language technology in business contexts.

In addition to rule-based semantics, as suggested in [15], it is necessary to build an organizational memory of terms and phrases, indexed using rule-based semantic annotation. Such vocabulary maintained as a Corporate Semantic Web (CSW) over an enterprise intranet, makes it convenient for the anyone in the enterprise to use it as a common repository. For the CP analysis, a section in CSW would contain specific indexes on product families, product features and services, warranty, claim procedures, etc.

3.2. Contextual Semantic Tagging

As said in [13], many semantic parsing systems depend on the information of Point-of-Speech (POS) tags that are abstracted, such as the Universal dependency tagset [19]. Such POS tags alone will not effectively classify the documents. In short, custom annotations addressed as *sem-tags* [13], refer to the task of assigning semantic class categories to the business-meaningful units in a sentence. *Sem-tags* will carry more information than universal POS tagging. For instance, in the car reviews data, "*Need driver seat memory and backup camera*" would represent seven tokens in POS tagging, while using a rule-based tagging, we would consider four token version "*Need driver_seat_memory and backup_camera*" suitable for semantic analysis. For the car reviews dataset, we describe below a sample of the semantic tagger we would recommend.

VEH (Vehicle)
- TYP type
- SUB subtype
- MAK make
- MOD model
- TRM trim
- MYR model year

ATTR (Physical Attributes)
- SIZ length / height / width
- WGT weight
- DRS number of doors
- COL color

ENG (Engine)
- EMK engine make
- ETP engine type (E.g.: 4 cyl.)
- POW horsepower / speed
- ACC torque / acceleration

PER (Performance)
- AWD all wheel drive
- HAN handling
- TRA transmission
- MPG fuel efficiency

INT (Interior)
- MAT material
- STO storage area
- SPA leg / head space
- VIS visibility
- STW steering wheel
- SUN sunroof
- LIT lighting
- DSH dashboard
- STY design / styling

SEAT (Seating)
- ROW location / number
- COM comfort
- ADJ adjustability
- MAT material
- SMR seat memory

SFTY (Safety & Security)
- BRK braking system
- AIR airbags
- SEN backup sensor / blindspot sensor
- LIT fog lights / high beams
- EMS emergency service
- LOK locking system

EXT (Exterior)
- BOD body
- AXL axles
- FRA frame
- TRK trunk
- WHL wheels / tires
- LIT lights
- MIR mirrors
- WIP wipers
- STY design / styling

ELE (Electronics)
- AUD audio system
- SCR display, video, camera
- BLU Bluetooth
- NAV navigation
- VOC hands free / voice command
- DAS dashboard / console
- REM remote start
- SEN sensors

Fig. 6. Custom annotations for car product reviews

3.3. Ontology

Ontologies [20] along with rules are used to represent the knowledge that will help with the specification of the concept pursued. Recently [21], systems are getting built for business data integration with ontologies to do interactive query systems. Ontologies have a structured framework that will guide the NLP process of analyzing CP. Ontologies for CP would require building domain knowledge and entities of interest into a traversable structure prior to launching the NLP pipeline. Let us show a possible ontology structure for the car reviews data [2]. An ontology for CP must cover car makes and models, interior and exterior features, engine types, to name a few. Car ontologies are used extensively for studying Advanced Driver Assistance Systems (ADAS) [22, 23]. For the example of the dataset described, a restricted car ontology context model based on [10] can be displayed as follows.

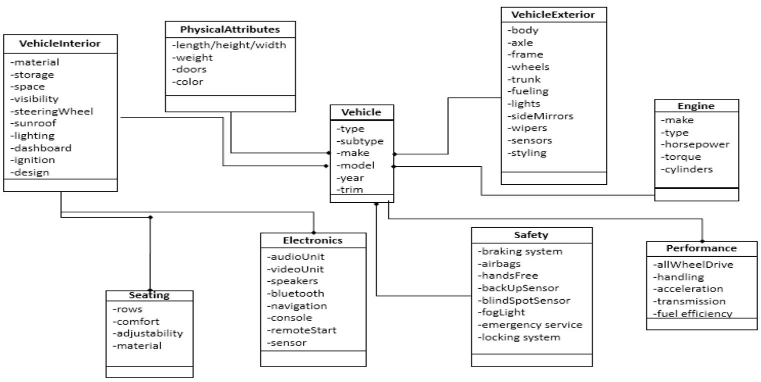


Fig. 7. Context model for product-based car ontology

The context model like the one designed in [21] starts with the vehicle design and its features. Its ontology plays a primary role in determining the interiors, exteriors, engine, and feature options details. Physical attributes are certainly a favorite attraction for any customer. Similarly, comfort, convenience, and drive performance factors play special role in determining the choice of a car. These could be manually added or captured in an automated way in the ontology. Some of the above language elements for CP can be developed using the deep learning. We discuss that in the next section.

4. Deep Learning concepts applying to CP

Deep Learning [24] is a multilayer neural network with multiple processing layers. It provides an in-depth analysis of new topics and facts, linking them to previously known concepts, or helps with formation of newer concepts for problem solving. It is broadly defined as feature representation learning according to [25]. It has been successfully applied to speech recognition [26], NLP [27] and image processing, just to quote a few. Before this concept was introduced, NLP modeling used to involve shallow architecture, with lots of human intervention in annotation, tagging and classification. Deep learning concepts of continuous-bag-of-words, skip-gram models etc., have revolutionized ever since they were first proposed [28].

Aim of the deep learning for CP analysis is to apply its methods to volumes of CP data to automatically generate context-based rich representation of required information. Many of the models, Convolutional Neural Nets, Recurrent Neural Nets, their special cases of Long Short-Term Memory (LSTM) and Gated Recurrent Units are applicable for CP. Though the machine translation uses deep learning [29] and is applicable to CP, we are restricting the concept discussion to CP data gathered in English texts only. However, machine translation is one area of NLP that has highly benefited through deep learning [29].

For CP data, Convolutional Neural Networks [30] help at a sentence level, identifying the rule-based sem-tagging. CNN will extract using the n-gram feature, create an informative semantic representation, that can be reviewed further with domain experts for conformity. They can also be built from a corpus of problem-specific Wikipedia documents, or product manuals etc. We can use word2vec for a large corpus of the documents and a pre-trained smaller test to build sentence classification. Performance of *word2vec* has been shown to improve if it is pre-trained for business contexts [31]. Convolution refers to the non-linear filters applied to sequence of words to produce a new feature. In CNN, many such convolution layers slide over word embeddings, pooled, and sub-sampled for efficiency. A caveat of CNN is optimization for long-distance dependencies, which may not be well-suited for CP data. The reason is, most of the customer views are expressed in short, sometimes incomplete sentences, or even cryptic grammars when it comes to mobile-app exchanges.

Recurrent Neural Networks (RNN) [27], on the other hand are sequential in their processing, are even better suitable for CP problems. In RNN, special cases of Long Short-Term Memory (LSTM) and the Gated Recurrent Units are applicable for CP specific business problems. RNN will help in sentence classification of customer views, and in turn, construct unsupervised classification of the customer reviews. The figure below explains clearly the framework of analyzing Customer Perception (CP) more efficiently using deep learning. Due to limitations of this article some of these concepts with examples will be explained in a later paper.

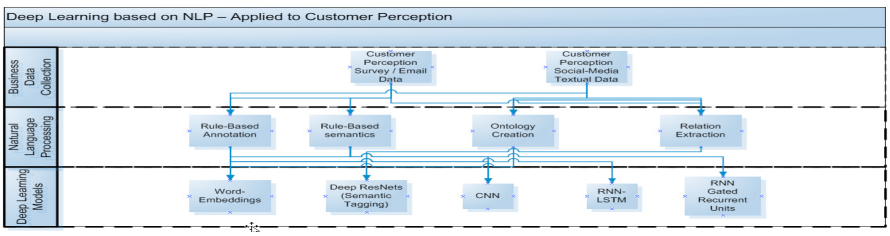


Fig. 8. Semantic tagging of documents from car reviews.

5. Data Analysis Approach

We explored the semantic tagging and business rule-based annotation for the car review dataset [2]. Since the work is elaborate, we present preliminary approach and results. As a first step, semantic annotation for the car specific comments are created as follows. This formal unified procedure is generic and applicable for all the features considered.

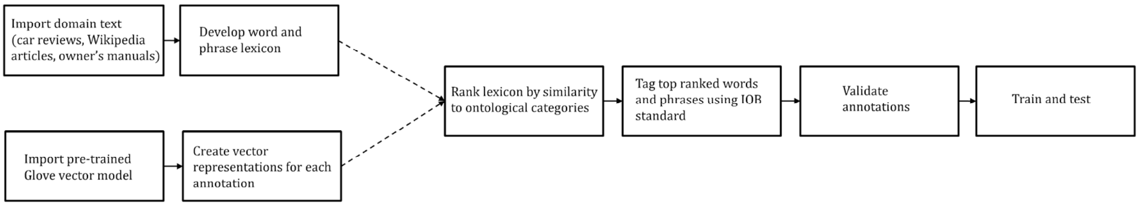


Fig. 9. Semantic tagging of documents from car reviews.

We import the domain specific text from car reviews, Wikipedia articles and owners' manuals and develop word and phrase lexicon. Using vector similarity of training document, semantic tagging following IOB (inside-outside-beginning) standards is done to annotate the document. We validate the annotations and apply manual corrections as necessary to curate our gold standard training set. Sample result of a tagged document for electronic features in a document is given in Fig.10.

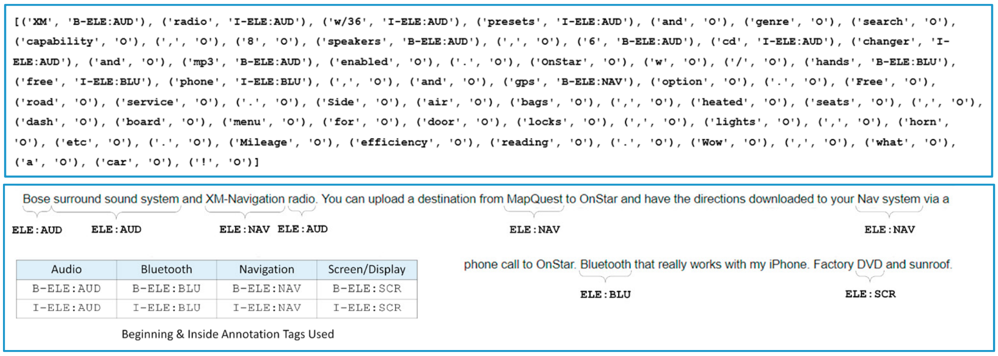


Fig. 10. Examples of semantic tagging and annotated documents.

Preliminary examples of the graph data model representation in Neo4j for Bluetooth and navigation electronics features are shown in Appendix A. This data model in conjunction with external data sources and the car reviews is used to automate much of the entity annotations prior to validation. Model training and validation on the car reviews dataset is ongoing as we expand the ontological pieces and capture more data.

6. Conclusion

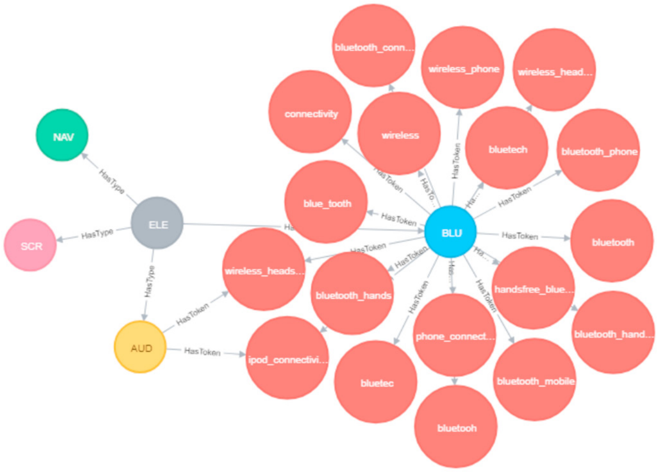
Customer Perception is gathered to understand product, brand or quality improvements, reduce claims, or gather information of customer views for new product introduction. In any of these cases, the customer views are shared through multiple mediums, and are subject to a contextual analysis in relation to the business questions considered. In this paper, we presented our recommendations for NLP technologies to be further explored for implementation of understanding customer perception. Primarily, rule-based annotation, semantic tagging and business ontology creation were explained through a dataset of car reviews [2]. These help to bridge the gap between the business expectation and the technical NLP analysis. In the end, the deep learning aspects most applicable to CP were explained. Future work aims to focus on complete corpus building and performing the deep learning CNN and RNN models mentioned here.

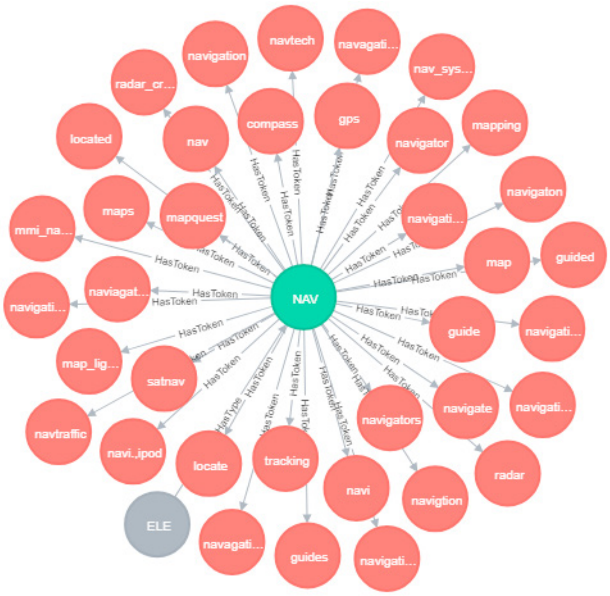
Acknowledgements

We would like to thank Caterpillar Analytics leaders, Chief Analytics Director Morgan Vawter, Analytics Division Manager Chaille Becker, and our supervisors, Daniel J. Organ and Fiona O'Lauglin for their support and encouragement.

Appendix A. Sample Graph Data Models for Electronics Features

Some of the graphical data representations of sub-features within electronics for the car dataset is shown below. Tags used are as provided in Figure 6.





References

[1] Conejo, F. (2013) "The Hidden Wealth of Customers: Realizing the Untapped Value of your Most Important Asset", (30)5 463-464.

[2] Ganesan, K. A., and ChengXiang Zhai, (2012) “Opinion-Based Entity Ranking”, *Information Retrieval.* 15 (2): 116-150.

[3] Dalgic, T. and Tulay Yeniceri (2013) *Customer-oriented marketing strategy: Theory and Practice*, Business Expert Press.

[4] Kureemun, K. and Robert Fantina (2011) *Your customers' perception of quality: what it means to your bottom line and how to control it*, New York: CRC Press.

[5] Narvaiza, L., David Ruiz de Olano, Tontxu Campos, and Iñigo Arroniz (2016) "When services are offered by small and medium retailers: what is the customer's perception?" *Universia Business Review* (**49**), 132-145.

[6] Sreejesh, S.; Juhi Gahlot Sarkar, Abhigyan Sarkar, Abdolreza Eshghi and M.R. Anusree (2018) "The impact of other customer perception on consumer-brand relationships", *Journal of Service Theory and Practice* (**28**)2 130-146.

[7] Ekekwe, Ndubuisi. (2010) “Mastering the Apple Game of Customer Perception.” *Harvard Business Review,* https://hbr.org/2010/10/mastering-the-apple-game-of-cu

[8] Song, J.H., and Candice R. Hollenbeck (2015) "The value of social presence in mobile communications", *The Service Industries Journal*, (**35)** 11-12, 611-632.

[9] Kaur, A., and Vishal Gupta (2013) “A survey on sentiment analysis and opinion mining techniques”, *Proc. of Journal of Emerging Technologies in Web Intelligence*, **5** (**4**) 349-354.

[10] Al-Moslmi, T., Nazlia Omar, Salwani Abdullah, and Mohammed Albared 2017) "Approaches to Cross-Domain Sentiment Analysis: A Systematic Literature Review", *IEEE Access,* **5**, 16173-16192.

[11] Nanli Z., Zou Ping, LI Weiguo, and Cheng Meng (2012) “Sentiment analysis: A literature review”, *Proc. of the International Symposium on Management of Technology (ISMOT), Hangzhou, IEEE,* 572-576.

[12] Gillick, D., Cliff Brunk, Oriol Vinyals, and Amarnag Subramanya (2016) “Multilingual language processing from bytes”, *Proc. of NAACL-HLT*, 1296–1306.

[13] Bjerva, J., Barbara Plank, and Johan Bos (2016) “Semantic tagging with deep residual networks”, *Proc. of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, Osaka, Japan, pp. 3531–3541.

[14] Luong, P-H., and R. Dieng-Kuntz (2007) “A rule-based approach for semantic annotation evolution”, *Computational Intelligence*, (**23**) 3, 320-338.

[15] Shalaby, W., and Wlodek Zadrozny (2017) “Mined Semantic Analysis: A New Concept Space Model for Semantic Representation of Textual Data”, *2017 IEEE International Conference on Big Data,* 2122-2131.

[16] Petasis, G., Frantz Vichot, Francis Wolinski, Georgios Paliouras, Vangelis Karkaletsis, and Constantine D. Spyropoulos (2001) “Using Machine Learning to Maintain Rule-based Named-Entity Recognition and Classification Systems”, *Proc. of the 39th Conference of Association for Computational Linguistics*, 418-425.

[17] Ponzetto, S.P., Roberto Navigli (2010) “Knowledge-rich Word Sense Disambiguation Rivaling Supervised Systems”, *Proc. of the 48th Annual Meeting of the Association for Computational Linguistics*, 1522–1531.

[18] Milne, D., and Ian H. Witten. (2008) “Learning to link with Wikipedia” *In Proc. of CIKM-08*, 509–518.

[19] Abzianidze, L. and Johan Bos (2017) “Towards Universal Semantic Tagging”, *CoRR* abs/1709.10381.

[20] Eiter T., Giovambattista Ianni, Axel Polleres, Roman Schindlauer, and Hans Tompits (2006) “Reasoning with Rules and Ontologies”, *Reasoning Web 2006, Lecture Notes in Computer Science*, (**4126**). Springer.

[21] Calvanese, D., Giuseppe De Giacomo, Domenico Lembo, Maurizio Lenzerini, Riccardo Rosati, and Gestionale Antonio Ruberti (2017) “Ontology-Based Data Access and Integration”, *Journal Encyclopaedia of Database Systems*, 1-7.

[22] Zhao, L., Ichise, Ryutaro Ichise, Seiichi Mita, and Yutaka Sasaki (2014) “An Ontology-Based Intelligent Speed Adaptation System for Autonomous Cars”, *Chapter in Graph Pattern Based RDF Data Compression,* 10.1007/978-3-319-15615-6_30.

[23] Feld, M. and Christian Muller (2011) “The Automotive Ontology: Managing Knowledge Inside the Vehicle and Sharing it Between Cars”, *Proc. of the 3rd Int. Conf. on Automotive User Interfaces and Interactive Vehicular Applications*, 79-86.

[24] Goodfellow, I., Yoshua Bengio and Aaron Courville (2016) *Deep Learning*, MIT Press.

[25] Deng, L. and Yu, D. (2014), "Deep Learning: Methods and Applications", *Foundations and Trends in Signal Processing*, (**7**): No. 3–4, 197-387.

[26] Rubi, C.R., (2015) “A Review: Speech Recognition with Deep Learning Methods”, *IJCSMC,* (**4**) **5** 1017–1024.

[27] Y. Goldberg (2017) *Neural Network Methods in Natural Language Processing*, Morgan and Claypool Publishers.

[28] Mikolov, T., Kai Chen, Greg Corrado, Jeffrey Dean (2013). “Efficient estimation of word representations in vector space”, *Proc. of the International Conference on Learning Representations*, 1-12.

[29] Jean, S., Cho, K., Memisevic, R. & Bengio, Y (2015). “On using very large target vocabulary for neural machine translation”, *Proc. ACL-IJCNLP* http://arxiv.org/abs/1412.2007.

[30] Hassan, A., and Ausif Mahmood (2018) “Convolutional Recurrent Deep Learning Model for Sentence Classification”, *IEEE ACCESS* **6** 13949-13957.

[31] Rong, Xin (2014). “word2vec parameter learning explained”, *arXiv preprint arXiv:1411.2738*.