A comparative study on vision, mission statement and goals of world-class automotive companies with text mining approach

Mohammad Mahdi Keshvari, Mohsen Shafiei Nikabadi

A comparative study on vision, mission statement and goals of world-class automotive companies with text mining approach

Mohammad Mahdi Keshvari and Mohsen Shafiei Nikabadi*

Industrial Management Department,
Binaoud Institute of Higher Education,
Mashhad, Iran
and
Industrial Management Department,
Semnan University,
Semnan, Iran
Email: Mohammad.Keshvari@gmail.com
Email: Shafiei@semnan.ac.ir
*Corresponding author

Abstract: This study explores the differences between the mission statements of Iranian and global automotive companies. This study identifies the gaps between the two by analysing the mission statements of fourteen international automakers and two Iranian automakers using Text Mining and *K*-means techniques. The analysis reveals that global companies are clustered into two groups based on the concepts of 'Custom' and 'Live'. In comparison, Iranian companies are clustered into two groups using the concepts of 'Benefit' and 'Abil'. The findings suggest that both foreign and Iranian companies tend to use vague concepts, but Iranian companies focus more on beneficence and company empowerment while prioritising the national market. Conversely, international companies prioritise enhancing customer satisfaction. This study highlights significant differences in the words used in Iranian and global automakers' vision and mission statements, emphasising the importance of understanding cultural nuances when creating these statements.

Keywords: data mining; text mining; clustering; automotive industry; *K*-means; world class.

Biographical notes: Mohammad Mahdi Keshvari holds his BSc degree in Computer Software Engineering and MSc degree in Industrial Management and Operations Research from the Industrial Management Department of Binaloud Institute of Higher Education. His research interests include information systems, operations research and data analysis.

Mohsen Shafiei Nikabadi is an Associate Professor of the Industrial Management Department at Semnan University, Semnan, Iran. He has been involved in multiple disciplinary researches and projects in the areas of knowledge management systems and performance optimisation in knowledge-based supply chains.

1 Introduction

The mission statement is a document that defines the organisation's Figure and unique goals: Why do we exist? What are our goals? What do we want to achieve? These are some of the critical questions that the mission statement intends to answer (Bart and Tabone, 2000). The vision for the organisation is like a realistic future, about to happen and attractive. Vision is an organisation's desired future and represents the situation it will achieve after successfully implementing its strategies. The vision statement expresses how the organisation wants to be seen in the future and its definition of success (Nanus, 1992).

Data mining involves various computational techniques constantly becoming critical approaches and effectively extracting new information from large data sets (Matthew and Karypis, 2004). Databases in text form are snowballing due to the proliferation of information available in electronic form. Today, most information in the industry, business and other organisations is stored electronically and in a text database (Imani, 2012). 80% of the stored data is textual (Gupta and Lehal, 2009). The increasing use of this information has motivated researchers to use new tools to discover patterns and gain knowledge. It is an interdisciplinary scientific textbook that includes computational linguistics, statistics, and machine learning. Exploring textual documents involves the linguistic and semantic analysis of the Text and its structure. As a result of this process, the hidden attributes of the Text have been revealed (Al-Azmi, 2013). Document clustering has been mentioned as one of the techniques in text mining in artificial intelligence. Document clustering as one of the machine learning methods without an observer, has many applications in various topics of natural language processing, such as information retrieval, automatic multi-text summarisation, etc. (Imani, 2012). This chapter will explicate the research problem, underscore its significance and need for investigation, articulate research objectives and inquiries, introduce innovative aspects of the study and delineate the methodology, including the rationale for using data analysis tools.

2 Literature review

Lagrand (2023) asserted the urge of automotive industries to respond to new innovations and the necessity to maintain a firm's competitive superiority By being able to detect emerging patterns and reflect that firms can more effectively manage their R&D portfolios by detecting some patterns via text mining. Kim and Trimi (2023) designed and delivered useful information for customers, which are crucial requirements of smart services using the results of data analysis, and introduced a methodology using morphological analysis and text mining for transforming data into information to support

smart services. Their methodology developed an application to the real-world case in smart services for vehicle operations management. Nikabadi and Varjovi (2020) reviewed 115 patents between 2000 and 2017 to study new areas in the aviation industry after selecting important factors for indexing and obtaining the weight of companies' indices using the AHP method and Chang's trigonometric fuzzy method. Allison (2019) reviewed 798 organisational vision and mission statements using the text analysis technique. King et al. (2019) studied changes in the Text of the Statements of Organisations in the most prominent offices in the USA. Lin et al. (2019) studied the Mission statements of 500 Chinese companies in various industries from the Fortune List with a linguistic approach. They analysed repetitive words with the NLPIR system and AntConc toolkit, which showed the differences and similarities between Chinese and non-Chinese companies. Zarei et al. (2018) investigated the scientific articles on the Competitive intelligence field through text mining in three steps and used k-means to cluster their findings. Shafiei Nikabadi and Karbasi Kheyr (2017) reviewed 53 companies from the list of top 500 companies of Fortune Magazine and 42 top Iranian companies in the field of the food industry, using the K-Means algorithm, clustering and text mining technique. Desmidt and Prinzie (2017) shown that the effectiveness and quality level of mission statements are related to companies' organisational performance. Nikabadi and Jafari (2016) found the characteristic of being green by examining the Text of the vision and mission statements of the top Iranian and international companies using text mining techniques. Nikabadi and Jafari (2016) analysed the vision and mission statement document, organisational mission and goals of 39 global insurance companies from the list of the top 500 companies worldwide and 18 Iranian insurance companies using text-mining and K-Means clustering and compared them with each other. David et al. (2014) analysed the Text of 9 changed companies' Mission Statements and created sentences that often orient the organisations towards customer orientation Instead of focusing on motivating internal managers. Rahimnia (2014) studied Top Iranian companies in Comparison with global companies; he analysed the vision and mission statement of top Iranian and international companies and looked for the green elements. Text analysis has recently been used in judicial and legal fields. It has been used in summarising texts that effectively identify characters, first names, verbs and most words that source voluminous documents (Al-Azmi, 2013). Another critical area in which text analysis has entered is social media. Given the powerful capabilities of text mining, Abdous and He (2011), Barbier and Liu (2011) and Pang and Lee (2008) believed that the use of text mining in social media data can lead to exciting discoveries in the field of human behaviour and interaction. In their survey, Nikabadi and Jafarian (2012) mentioned IKCO and Saipa as two prominent automotive companies in Iran. AlShameri et al. (2012) clustered 772 companies in the Fortune 500 using vision and mission statement documents. They demonstrated how these clusters could be explained by globalisation, innovation and strategy features. Breznik and Dermol (2012) examined the mission statement documents of prominent Slovenian companies. Dimensions extracted from word network analysis were presented in four concepts: 'quality of life', 'growth', 'cooperation and innovation' and 'customer'. Hanafizadeh and Nikabadi (2011) studied Iranian automotive companies and found only Ikco and Saipa possess holding and chain structures.

Li et al. (2012) also used text mining in the fields of medical sciences, educational fields (Hung, 2012; Abdous and He, 2011) and business (Ingvaldsen and Gulla, 2012). Fuller et al. (2011) used deception detection in real-world data with data mining

techniques. Nasehifar and Pourhosseini (2008), in a study named 'Comparative study of the components of the mission statement of local IT companies with foreign companies,' have examined the mission statement of Iranian and foreign IT companies in terms of similar components of the usual criteria in the mission statement. Francis (2006) and Weiss et al. (2010) presented a detailed article on using clustering and other unsupervised techniques in text mining. Textbook research in health and medicine has been used to link several hundred medical records and find relationships between symptoms and prescriptions (Smith, 2010). Williams (2008) compared the mission statements of top companies (listed in the Fortune 1000 list) with low-efficiency companies using content analysis and text analysis methods. Kolyshkina and Van Rooyen (2005) presented their analysis results using text mining in the automotive Complaint report database. Davi et al. (2005) demonstrated the use of data mining in verifying key steps for analysing unstructured texts. Hossain (2004) examined the mission statement of 15 Bangladeshi companies based on the David model and concluded that these companies do not pay attention to the basic features of this document in their mission statement. Many studies have been done on text mining and its applications. A case study on the use of text mining in patent analysis in 2003 compared to PackMOLE. The authors showed that the software for the manual patent analysis technique has more advantages (Fattori et al., 2003). Witten et al. (2004) used text mining techniques to extract metadata from documents in a library and to help enrich records by highlighting appropriate items in the Text. They found that text mining could add value to digital library documents and the library user experience. In the field of clustering, archived documents are also used in several pre-defined clusters and semantic classifications (Neto et al., 2000). Pearce and David (1987) reviewed Fortune Magazine's 500 selected companies in terms of high and low performance. Pearce and David (1987) suggested they could not find any reported studies detailing the composition of business mission statements, and they only found a limited number of conceptual articles suggesting desired component characteristics of mission statements. Yet Cortes and Duenas (2022) somehow admitted that this and other gaps still exist.

Looking at the history of international automotive companies that have been initialised simultaneously with Iranian automotive companies and comparing their growth rates with each other, one can get a glimpse of the enormous gap created between them. In addition to financial problems and technological issues, Other issues, such as the company's organisational goals as a strategic policy, should also be re-examined to ensure the accuracy of the plans in Iranian automotive companies. Therefore, studying the text of large companies' mission statements, vision and organisational goals is imperative to finding their most essential components and orientation. For this purpose, we need a tool for analysing texts to review documents related to organisational goals (vision and mission statement) and classify them.

3   Research method and steps

This research seeks to find the flaws and discrepancies in the vision and mission statements of Iranian versus international automotive companies utilising text-mining techniques to analyse and subsequently cluster them in a series of steps.

3.1   Creating a database

Companies are selected from the list on the Business Insider (2018) website. The data, which contains the vision, mission statement, and goal of the world's top international automotive companies, is gathered from the official websites of BMW, GEELY, Volkswagen, DAIMLER, FCA, PSA, Hyundai, General Motors, Renault, Nissan, TATA, TOYOTA, Ford and HONDA. The following Iranian companies are IKCO and Saipa.

3.2   Pre-processing

To analyse the content of texts, we must first perform activities on them to be prepared for review and processing. We want to turn unstructured texts into structured data that allows us to make textual analysis operations on them (Weiss et al., 2010). In a study of data science tools by Senan and Tasci (2024), RapidMiner was used as one of the four data processing tools for performing the operations required for data mining. We have used the text mining plugin in RapidMiner to perform text mining operations.

3.3   Text processing

In the present study, we used the TF-IDF word vector output, showing us the importance of a word in a text. The document processing plugin enables us to prune so-called texts. In this study, we selected the maximum amount as a relatively large number and the minimum repetition of each word in all documents to one in order to observe all-important candidate words as much as possible.

3.4   Tokenise

This section decomposes symbols, phrases, words or other meaningful elements. At this point, we aim to discover the words in the Text. Initially, textual data should be only a snippet of characters or a text description, while we need assertions in the data collection phase. So, we need a parser for the word separation operation (Weiss et al., 2010).

3.5   Transform cases

For this part, we eliminate the sensitivity to uppercase and lowercase letters. In this way, all the characters become uppercase or lowercase letters, and the user can choose all uppercase or lowercase letters (Verma et al., 2014).

3.6   Filter stop words

This filter is used to clear very general words that have a negative impact on the text mining process. Items such as pronouns, conjunctions and prepositions can be recognised as ineffective words. Since many texts deal with such cases, this filter can delete all these categories of words.

3.7   Stem operator

To reduce the number of words and avoid examining different word types, we use the root finder function to count words derived from a particular root only once. In this way, the terms become shorter and closer to their roots.

For example, efficiencies, efficacious, efficiency, efficient are all from the root efficacy, which should be referred to their short roots since we are not dealing with slang in human relations and we are also working more with book texts in the form of a vision and mission statement, we will not face such problems. So, using the root function will continue to be helpful because the suffixes of one verb and the other forms of a word all go back to the same root (Verma et al., 2014).

3.8   Filter tokens by length

This function is to specify the size of the parts (word, phrase, sentence, etc.) that the program specifies (Verma et al., 2014). We can determine the maximum and minimum lengths of the words used in this study.

3.9   Generate N-Grams

This filter specifies a sequence of words or roots of length *N* that are repeated in the Text. Sometimes, we come across phrases or comments in the Text that follow each other and create a specific semantic load. So, using this word filter, we identify and examine the roots of words used repeatedly in the Text. In this filter, we can specify the maximum number of following terms (Weiss et al., 2010).

3.10   Formation of TF-IDF matrix

This function is usually used to retrieve data from TF-IDF weight values. TF-IDF consists of 2 parts. The first component calculates the normalised value of word repetition in the Text, which is called TF. This value is equal to the number of times a word appears in a text, divided by the total number of words in that Text. TF indicates the frequency of an expression in the Text. A comment may be divided more often into long texts than short texts. Obviously, the number of words in the Text is to be normalised. The second part is the inverse of the number of documents, called the IDF. This value is calculated as the logarithm of the total number of copies in the Text set divided by the number of records in which a particular word appears. It measures the importance of a comment. Finally, the final TF-IDF value for each term will be obtained by multiplying these two values. Combining these values for different words and documents in the relevant matrix will be brought (Wu et al., 2008).

3.11   Clustering

After taking the data from the output of the previous function, the TF-IDF matrix is available, which can be used for clustering in this section. Now, we have a numeric matrix that we can use to compare any document. Consider it as a vector of words that determines the order of words in all copies. Hence, we will be able to start the clustering

operation. Clustering is related to grouping components so that the members have the most differences compared to other clusters, and similar sets have the most similarities.

These differences and similarities are determined by the distance of objects from each other (Han et al., 2011). By focusing on this feature, we can identify similar texts related to the vision and mission statement of the companies under review and have a cluster of each. We can assign each object to only one set with the *K*-means algorithm.

3.12 Determine the value of K

K-means clustering works by assigning points as cluster centres and moving them to achieve the best clustering. We use a random value for *K*, the number of clusters' centres, to start the iteration loop. Then, we assign the data point to the nearest declared centres. Finally, we change the location of the centres so that their values are equal to the average values of the data. This must be repeated to converge to the best approximation. The performance of *K*-means depends on the correct choice of centres. Therefore, the number of cycle repetitions is approximated to achieve better results. We obtained 2 clusters for both Iranian and foreign companies. Also, to select more appropriate initial values for determining the centres of the clusters, we used the Arthur and Vassilvitskii (2007) method to determine the best deals for starting the matters related to the cluster centres. We ranked the derived roots according to the given values in the next step. The output shows the rate of word occurrence in all extracted texts, and the average TF-IDF values of every single seed make a difference in document text (Klinkenberg, 2013).

3.13 Semantic review of the text and the expert's approval

In the next step, we did not limit ourselves to the software output because it is possible that the selected root in the software would be regardless of any meaning and concepts. Hence, we find the desired sources in the Text of each document by considering the importance of each root in the Text; we can see that interpretations related to the roots are drawn out. Finally, according to the use in the Text and the semantic load, some of the origins are rejected and removed from the list of words.

In the next step, we expose the obtained roots to the automotive industry experts to identify and confirm several more important roots based on essential characteristics in the automotive industry. As a result, the roots are carefully selected and finally approved by experts. In this way, each cluster indicates the most important roots in terms of organisational perspective and goals (Shafiei Nikabadi and Karbasi Kheyr, 2017).

4   Findings

Based on the previously mentioned criteria, the clustering stage results from the texts of the mentioned companies placed in individual clusters. Figure 1 shows the claustration of international companies. This cluster is based on the TF-IDF values of each company document's most frequent and essential words. Top global automotive companies are classified into two groups. Each set has companies that are most similar in their vision, mission statement and organisational goals.

Figure 1   Clustering of international car companies (see online version for colours)

Clusters	Company
Cluster 1	BMW – GEELY – Volkswagen - DAIMLER – FCA – PSA
Cluster 2	Hyundai – General Motors – Renault – Nissan –TATA – TOYOTA – Ford – HONDA

Figure 2 shows the clusters of Iranian automotive companies. The text mining plugin in RapidMiner has been used as the tool for text mining and then clustering them with the *k*-mean algorithm. Wordle software has been long used for the graphical presentation of clusters and concepts. Wordle software uses textual data provided by the researcher and generates 'word clouds.' Clouds highlight more frequent words in the Text and display them in various fonts, layouts and colours (AlShamri et al., 2012). We use another cloud service that precisely states what cloud-based WordLe software does to display words, which is called WordCloud, and we use this service to show our data graphically.

Figure 2   Clustering of Iranian car companies (see online version for colours)

Clusters	Company
Cluster 1	IKCO
Cluster 2	SAIPA

Figure 1 lists the names of six international automotive companies in the first cluster and the names of eight global automotive companies in the second cluster.

Figure 2 shows the names of Iranian automotive companies in the first and second clusters.

4.1   Data vector output

After completing the data gathering and processing steps, the texts are converted into words and presented in a tabular structure where each record represents one of the documents. Each column represents one of the words. Figure 3 provides weighted examples of TF-IDF vectors and frequently used words in international automotive company documents.

Figure 3   TF-IDF vectors and frequently used word examples in the documents of international automotive companies (see online version for colours)

Row	World	Custom	Mobil	lead	People	Live
1	0.168	0.219	0.258	0.284	0.203	0.200
2	0.145	0.111	0.177	0.136	0.179	0.153
3	0.093	0.103	0.118	0.125	0.137	0.147

Figure 4 presents the TF-IDF vector and the most frequent words in the documents of international car companies in actual events and the number of records.

Figure 4   TF-IDF vector and repetitive words in the documents of international car companies
            (see online version for colours)

Word	Total	Document
World	11.0	9.0
Custom	9.0	8.0
Mobil	8.0	6.0
Lead	7.0	6.0
People	7.0	5.0
Live	6.0	5.0

Figure 5 shows the TF-IDF vectors and the most frequently used words in the documents of Iranian car companies by weight.

Figure 5   TF-IDF vectors and repetitive words in the documents of Iranian car companies (see
            online version for colours)

Row	market	active	product	abil	benefit	company
1	0.058	0.087	0.087	0.0	0.277	0.277
2	0.213	0.159	0.159	0.255	0.0	0.0

Figure 6 provides the TF-IDF vector and the most frequently used words in the documents of Iranian automotive companies in the form of the entire event and the number of records.

Figure 6   TF-IDF vectors and repetitive words in the documents of Iranian car companies
            (see online version for colours)

Word	Total	Document
Market	5	2
Active	3	2
Product	3	2
Abil	2	1
Benefit	2	1
Company	2	1

4.2   Determining the relationship between clustering and labels

The output of the data processing stage is in the form of a list of words and a vector of documents. The word list is based on the number of documents in which the utterance occurred and the repetition of the word in all documents. Owing to the fact that some words have happened in a small number of documents and, on the other hand, some have been publicly repeated in most documents, the words that have been repeated in more

than three documents have been included in the documents related to international companies. In Iranian companies, Words have been studied by repeating more than 1 document. In the present study, the *K*-mean clustering method has been used. In this method, the number of clusters in the range of 1 to 8 was finally considered as 2 for a set of international companies (see Figure 7) and 2 for Iranian companies (see Figure 8). Figure 7 shows the weighted average of repetitive words in two clusters.

Figure 7   Clustering of international automotive companies (see online version for colours)

Word	Cluster 1	Cluster 2
World	0.052	0.063
Custom	0.080	0.043
Mobil	0.080	0.047
Lead	0.021	0.081
People	0.021	0.081
Live	0	0.091

Figure 8 shows the weighted average of frequently used words of Iranian automotive companies in two clusters.

Figure 8   Clustering of Iranian car companies (see online version for colours)

Word	Cluster 1	Cluster 2
Market	0.058	0.213
Active	0.087	0.159
Product	0.087	0.159
Abil	0	0.255
Benefit	0.277	0
Company	0.277	0

In the following, the relationship between words and clusters is shown in a diagram, which can be used to extract the concepts and words that create the most separation between clusters and be used as a cluster label. Figure 9 is a diagram of international car companies. Based on this diagram, the word Custom separates the first cluster from the other cluster and is selected as the cluster's label. The word Live is chosen as the cluster label in the second cluster.

Figure 9 Diagram of the relationship between words and clusters in international automotive companies (see online version for colours)
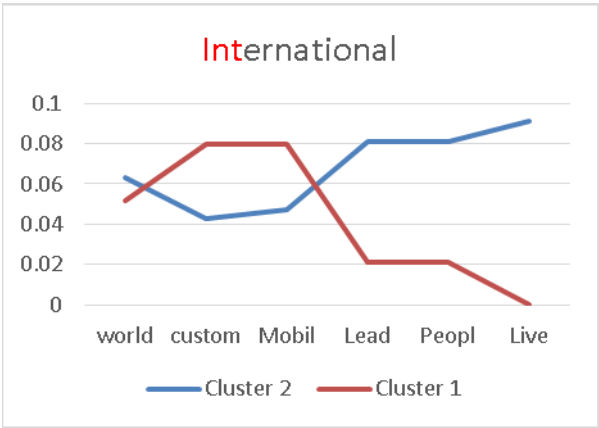

Figure 10 is a diagram of Iranian automotive companies. Based on the above argument, the word Benefit has been selected as the label in the first cluster, and in the second cluster, the word Abil has been chosen as the cluster label.

Figure 10 Diagram of the relationship between words and clusters in Iranian automotive companies (see online version for colours)
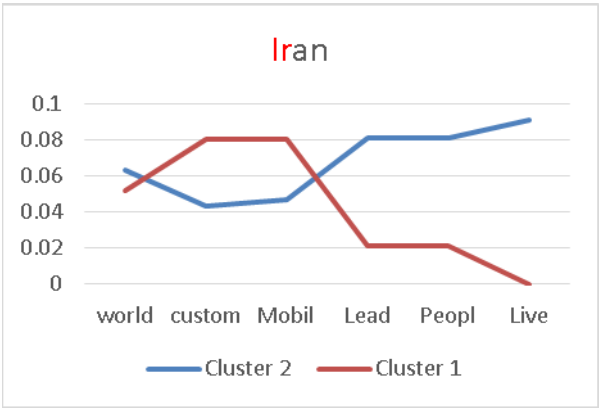

Figure 11 shows the labels of the first six international automotive companies, including BMW, GEELY, Volkswagen, DAIMLER, FCA and PSA. This Figure shows the most frequently used words in this cluster.

Figure 11 Labels of International automotive companies in cluster 1 (see online version for colours)

Figure 12 shows the second cluster of international automotive companies, including eight companies: Hyundai, General Motors, Renault, Nissan, TATA, TOYOTA, Ford and HONDA. This Figure shows the most frequently used words in this cluster.

Figure 12 Label of international automotive companies in cluster 2 (see online version for colours)

Figure 13 shows the labels of the first cluster of Iranian automotive companies and their names. In this Figure, the most frequently used words are displayed.

Figure 13 Labels of Iranian automotive companies in cluster 1 (see online version for colours)

Figure 14 shows the labels of the second cluster of an Iranian automotive company and its name: Saipa. In this Figure, the most frequently used words are displayed.

Figure 14 Labels of international automotive companies in cluster 2 (see online version for colours)

6 Discussion and conclusion

The current chapter aims to compare and analyse the obtained clusters and corresponding results.

6.1 Analysis and comparison of findings in clusters of Iranian and foreign companies

Upon examining the vision, mission statements and goals of Iranian and international automakers, the authors identified thought-provoking concepts in the labels of the obtained clusters, which are presented in Figure 15. By analysing this Figure and separately studying the companies' vision, organisational mission and goals based on the extracted concepts from the clusters, the authors obtained the following results.

Figure 15 Cluster labels of automotive companies (see online version for colours)

Cluster Labels (International Automotive Companies)	Cluster Labels (Iranian Automotive Companies)
Customer Orientation	Profitability
people life	ability

6.1.1 Comparison of Cluster 1 (international companies) and Cluster 1 (Iranian companies)

Global companies in the first cluster have used customer orientation specifically in 3 cases. In other cases, the word customer relationship has been precisely oriented in the Text of these documents, and concepts such as customer orientation and communication have been used. In contrast, Iranian companies in the first cluster have used customer satisfaction as a secondary complication, indicating a lack of planning and particular communication for customer-oriented activities. More focus on customer service response in terms of after-sales service and creating a more robust and more effective customer relationship network can be helpful for this cluster of Iranian automakers to get closer to international automotive companies.

6.1.2 Comparison of Cluster 1 (international companies) and Cluster 2 (Iranian companies)

International companies belonging to Cluster 1 have used the concept of customer orientation alongside terms such as global in their documents publicly and privately, and concepts such as customer relationship and customer satisfaction are seen in their papers. In cluster 2 (Iranian companies), the root of the word customer is brought next to the Iranian customer and localisation. It considers paying attention to customer needs depends on local conditions and tastes but then expands to words such as 'reach global', 'global markets', and 'to reach the market'. The Iranian cluster is trying to reach the international market by tunnelling from national tastes and adapting to national standards, which is noteworthy. Iranian companies should seek and create solutions that suit better mechanisms to understand the needs and preferences of their customers at international and global levels and fulfil these needs to align themselves with the market.

6.1.3   Comparison of Cluster 2 (international companies) and Cluster 1 (Iranian companies)

Global companies have used the word 'life' specifically in 6 cases. In other cases, the words customer, people's lives, souls and people have been precisely oriented in the Text of the documents, and concepts such as the impact of their products on people's life quality have been used. On the other hand, Iranian companies generally do not go beyond the general use of the word customer satisfaction and do not seem to be particularly concerned. It appears that Iranian companies look only at their products only as good to sell and ignore them as a product that has a remarkable impact on customers' quality-of-life. Iranian automotive companies refer to customer satisfaction only limited to sales and after-sales service. They do not go beyond that, which requires a review of the strategy and approach of Iranian automotive companies to further align with international automotive companies.

6.1.4   Comparison of Cluster 2 (international companies) and Cluster 2 (Iranian companies)

Global companies belonging to cluster 2 have used the concept of paying attention to the impact of their products on people's life quality in their documents, but in cluster 2 (Iranian companies), except for the general use of the term Iranian market and customer, other cases which specifically states their views and concerns were not observed. Iranian companies do not seem to pay attention to their product impacts on the life quality of their customers and are focused on manufacturing and selling their products without getting feedback on their products' effects on their customers' lives. This view may have been created due to Iran's lack of a competitive market due to the current economic and political situation and regulations.

6.1.5   Cluster Survey 1 (International Companies)

The study of this cluster shows that customer orientation is explained in the documents of these companies and is described in the form of words such as customer retention, communication, attending to customer needs and customer relationship. Companies have related the meaning of 'value' to their customers.

6.1.6   Cluster Survey 2 (International Companies)

Examination of this cluster shows that the word 'life' is explained in the documents of these companies and is written in the form of words such as customer, people's lives, lives and people. In this way, companies have shown their customers the theme of the phrase life by showing the impact of their products on people's life quality to their customers.

6.1.7   Cluster Survey 1 (Iranian Companies)

An examination of this cluster, which has been labelled with the word profitability, shows that focusing on financial benefits, company growth, investing and producing added value in the resource sector is the primary goal of this cluster.

6.1.8 Cluster Survey 2 (Iranian companies)

The study of this cluster shows that the companies presented in this cluster have mentioned the word capability to enter the global market as an approach in their vision, mission statement, and goals. However, without any enlightenment and definition of how to pay attention to the Iranian market and make changes in localisation direction. How does it end up achieving world-class needs and not international ones.

7 Conclusion and future works

It is concluded that international automotive companies have used precise and defined words in their vision and mission statement documents and clearly explained their customer approach. It is inferred that these companies are looking for world-class activities and must be accurately oriented, value-based and have an economic orientation in their vision and mission statements. The impact on stakeholders' lifestyles and life quality is more important to them than profitability. In contrast, Iranian companies, except for expressing some concepts of profitability and empowerment, an integrated part of most companies' vision and mission statement, did not provide details on other ideas of customer orientation and humanism. They used localisation and customer satisfaction as their general meaning. Undoubtedly, the viability of companies at present at international levels requires reconsideration of the vision and mission statement for these companies and adaptation to global perspectives.

Examining the vision and mission statements of the Iranian automotive companies, they have paid very little attention to customer needs. They have focused most of their attention on production, creating added value and finally expanding to foreign markets. Still, the locking point in this approach is that international car companies seem to have a shift in fulfilling the demands of today's customers by raising their quality-of-life, which can be interpreted in various ways, such as safety and well-being while using those products, as shown in the documents of international companies. However, Iranian car companies have not paid attention to this matter. They have focused on their company's profitability, production, empowerment and finding a way to enter international markets through their resolutions. The organisations' approach to their current vision and mission statement documents for the automotive industry was unclear. This mindset, arguably may be seen in these companies' rapid unjust manipulation of the Iranian automotive market. Also, influential and essential factors explaining these documents' theories are questionable. The researcher was not successful in finding research or articles to find an answer to this question. Other Iranian automotive companies are also active in the Iran market. Still, due to a lack of access to genuine documents and not receiving any response from these companies indicating any registered national car brand, we were limited to examining two well-known automotive companies in Iran with national registration certificates.

7.1 Limitations and concerns of the study

The researcher was unable to find any research or articles to answer the question of how an organisation approaches the creation of a 'Vision' document and organisational 'Mission Statement' for the automotive industry. Additionally, the investigation of the

two largest automobile companies in Iran, which have products with national registration certificates, is limited to them, mainly because of the lack of response from other automobile companies in the domestic industries of the country or the absence of documents indicating that they have a domestic patented national brand. The FCA and PSA companies merged as a new company named Stellantis, but since no reference was found to rely on for the new order, it is concluded that it is better to stick with the referenceable name lists.

7.2   Discussion and future work

There are many models in the field of organisational mission and practical concepts in its formulation, among which we can refer to David's strategic model, strategic management, which examines the Text of a company's organisational mission and compares the concepts of text mining with the concepts presented in David's (2007) model. We also suggest analysing the alignment degree of the company's mission statement and the model above. In addition, it is recommended to compare the most important dimensions of the clustering in the present study with the dimensions of the balanced scorecard of organisations and their performance uses in another study. Examining and analysing the report of automaker complaints is very effective in changing the behaviour and strategy of automotive companies. This can be achieved by conducting continuous research and data mining, particularly text mining. It is also suggested to examine the history of changes in companies' vision and mission statements to see how and why changes happened in vision and mission statement documents and goals of companies. It is interesting to monitor the vision and mission statements of the companies every 5 to 10 years to watch the changes in the companies' insights and compare them together.

References

Abdous, M. and He, W. (2011) 'Using text mining to uncover student's technology related problems in live video streaming', *British Journal of Educational Technology*, Vol. 40, No. 5, pp.40–49.

Al-Azmi, A. (2013) 'Data, Text, and web mining for business intelligence: a survey', *International Journal of Data Mining and Knowledge Management Process*, Vol. 3, No. 2. Doi: 10.5121/ijdkp.2013.3201.

Allison, J. (2019) 'Mission statements and vision statements: examining the relationship toward performance outcomes', *Global Journal of Management and Marketing*, Vol. 3, No. 1. pp.1–21.

Alshameri, F., Greene, G.R. and Srivastava, M. (2012) 'Categorizing top fortune company mission and vision statements via text mining', *International Journal of Management and Information Systems (Online)*, Vol. 16, No. 3, pp.227–238.

Arthur, D. and Vassilvitskii, S. (2007) 'K-means++: the advantages of careful seeding', *Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms*, Society for Industrial and Applied Mathematics, pp.1027–1035.

Barbier, G. and Liu, H. (2011) 'Data mining in social media', Aggarwal, C. (Ed.): *Social Network Data Analytics*, pp.327–352.

Bart, C.K. and Tabone, J.C. (2000) 'Mission statements in Canadian not-for-profit hospitals: does process matter?', *Health Care Management Review*, Vol. 25, No. 2, pp.45–63.

Breznik, K. and Dermol, V. (2012) 'Identifying the mission statement's clusters in Slovenian economy', *British Journal of Educational Technology*, TIIM Lublin, Poland, Vol. 43, No. 1, pp.5–16.

Business Insider (2018) *Business Insider*. Available online at: https://www.businessinsider.com/biggest-car-companies-in-the-world-details-2018-2

Cortes, J. and Duenas, J. (2022) 'What is the message of mission statements?', *Academy of Management Proceedings*. Doi: 10.5465/AMBPP.2022.10083abstract.

Davi, A., Haughton, D., Nasr, N., Shah, G., Skaletsky, M. and Spack, R. (2005) 'A review of two text-mining packages: Sas text mining and wordstat', *The American Statistician*, Vol. 59, No. 1, pp.89–103. Doi: 10.1198/000313005X22987.

David, M.E., David, F.R. and David, F.R. (2014) 'Mission statement theory and practice: a content analysis and new direction', *International Journal of Business, Marketing, and Decision Sciences*, Vol. 7, No. 1, pp.95–110.

Desmidt, S. and Prinzie, A.A. (2009) 'The effectiveness of mission statements: an explorative analysis from a communication perspective', *Academy of Management Proceedings*, 1–6. Doi: 10.5465/ambpp.2009.44257947.

Fattori, M., Pedrazzi, G. and Turra, R. (2003) 'Text mining applied to patent mapping: a practical business case', *World Patent Information*, Vol. 25, No. 4, pp.335–342.

Francis, L.A. (2006) 'Taming text: an introduction to text mining', *Casualty Actuarial Society Forum*, pp.51–88.

Fuller, C., Biros, D. and Delen, D. (2011) 'An investigation of data and text mining methods for real world deception detection', *Expert Systems with Applications*, Vol. 38, No. 7, pp.8392–8398. Doi: 10.1016/j.eswa.2011.01.032.

Gupta, V. and Lehal, G. (2009) 'A survey of text mining techniques and applications', *Journal of Emerging Technologies in Web Intelligence*, Vol. 1. Doi: 10.4304/jetwi.1.1.60-76.

Han, J., Pei, J. and Kamber, M. (2011) *Data Mining: Concepts and Techniques*, 3rd ed., Elsevier, pp.429–445.

Hanafizadeh, P. and Nikabadi, M.S. (2011) 'Framework for selecting an appropriate e-business model in managerial holding companies: case study: Iran Khodro', *Journal of Enterprise Information Management*, Vol. 24, No. 3, pp.237–267. Doi: 10.1108/17410391111122844.

Hossain, D.M. (2004) 'An analysis of the mission statements of selected Bangladeshi Companies [Bangladesh]', *Journal of Business Research*, Vol. 6, pp.83–91.

Hung, J. (2012) 'Trends of e-learning research from 2000 to 2008: use of text mining and bibliometrics', *British Journal of Educational Technology*, Vol. 43, No. 1, pp.5–16.

Imani, M. (2012) *Clustering of Persian Texts*, Master Thesis, University of Science and Technology (Persian), Tehran.

Ingvaldsen, J.E. and Gulla, J.A. (2012) 'Industrial application of semantic process mining', *Enterprise Information Systems*, Vol. 6, No. 2, pp.139–163. Doi: 10.1080/17517575.2011.593103.

Kim, M. and Trimi, S. (2023) 'Transforming data into information for smart services: integration of morphological analysis and text mining', *Service Business*, Vol. 17, pp.257–280. Doi: 10.1007/s11628-023-00526-y.

King, D.L., Case, C.J. and Carney, J.R. (2019) 'Evolving mission statements: a fifteen year comparison', *Global Journal of Business Pedagogy*, Vol. 3, No. 1. Doi: 10.47177/gjbp.03.01.2019.06

Klinkenberg, R. (Ed.) (2013) *RapidMiner: Data Mining Use Cases and Business Analytics Applications*, Chapman and Hall/CRC, pp.242–261.

Kolyshkina, I. and Van Rooyen, M.M. (2005) 'Text mining for insurance claim cost prediction', Presented at *the 15th General Insurance Seminar Institute of the Actuaries of Australia*, pp.1–24.

Lagrand, B. (2023) *Exploring the advancements and future research directions of artificial neural networks: a text mining approach*, Bachelor's thesis, University of Twente. Available online at: http://essay.utwente.nl/95448/1/Lagrand_BA_BMS.pdf

Li, L., Ge, R.L., Zhou, S.M. and Valerdi, R. (2012) 'Guest editorial: integrated healthcare information systems', *IEEE Transactions on Information Technology in Biomedicine*, Vol. 16, No. 4, pp.515–517. Doi: 10.1109/TITB.2012.2198317 PMID:22760931.

Lin, Q., Huang, Y., Zhu, R. and Zhang, Y. (2019) 'Comparative analysis of mission statements of Chinese and American Fortune 500 companies: a study from the perspective of linguistics', *Sustainability*, Vol. 11. Doi: 10.3390/su11184905.

Matthew, R. and Karypis, G. (2004) *Soft Clustering Criterion Functions for Partitional Document Clustering*, Technical Report, Department of Computer Science and Engineering, Universit og Minnesota.

Mullane, J.V. (2010) 'The mission statement is a strategic tool: when used properly', *Management Decision*, Vol. 40, No. 5, pp.448–455.

Nanus, B. (1992) *Visionary Leadership: Creating a Compelling Sense of Direction for Your Organization*, Jossey-Bass Inc., 350 Sansome Street, San Francisco, CA 94104-1310.

Nasehifar, F.V. and Pourhosseini, S. (2008) 'A comparative study of mission statement components of IT companies of domestic and foreign countries', *Management Knowledge*, Vol. 21, No. 80, pp.123–142. (In Persain).

Neto, J., Santos, A., Kaestner, C. and Freitas, A. (2000) 'Document clustering and text summarization', *Proceeding of the Fourth International Conference Practical Applications of Knowledge Discovery and Data Mining*, London, UK.

Nikabadi, M. and Jafarian, A. (2012) 'Framework for selecting an appropriate e-business model in SMEs', *International Journal of E-Business Development (IJED)*, pp.86–96.

Nikabadi, M.S. and Jafari, O.A.Z. (2016) 'Clustering and comparative study of vision document, organizational mission, and objectives of international and domestic insurance companies with text mining approach', *Journal of Insurance*, pp.1–24. (Persian)

Nikabadi, M.S. and Varjovi, H. (2020) 'Identifying new business areas in air transport industry using patent information', *International Journal of Business Innovation and Research*, Vol. 1, No. 1.

Pang, B. and Lee, L. (2008) 'Opinion mining and sentiment analysis', *Foundations and Trends® in Information Retrieval*, Vol. 2, Nos. 1/2, pp.1–135. Doi: 10.1561/1500000011.

Pearce, J.A. and David, F. (1987) 'Corporate mission statements: the bottom line', *The Academy of Management Executive*, Vol. 1, pp.109–115. Doi: 10.5465/ame.1987.4275821.

Pearce, J.A. and David, F. (1987) 'Corporate mission statements: the bottom line', *The Academy of Management Executive*, Vol. 1, No. 2, pp.109–116.

Rahimnia, F. (2014) 'Green organizational mission and vision, a case study: top Iranian companies compared with top global companies', *Third International Conference of Behavioral Science*, pp.58–64.

Shafiei Nikabadi, M. and Karbasi Kheyr, A. (2017) 'A comparative analysis of mission and vision statements of top fortune companies and Iranian companies in food industry using text mining and clustering', *Futurology of Management*, Vol. 28, No. 4 (111), pp.29–52. (Persian)

Smith, D. (2010) *Using Data and Text Mining to Drive Innovation*, PhUSE UK.

Verma, T., Renu, R. and Gaur, D. (2014) 'Tokenization and filtering process in RapidMiner', *International Journal of Applied Information Systems*, Vol. 7, No. 2, pp.16–18.

Weiss, S.M., Indurkhya, N. and Zhang, T. (2010) *Fundamentals of Predictive Text Mining*, Vol. 41, Springer. Doi: 10.1007/978-1-84996-226-1.

Williams, L.S. (2008) 'The mission statement: a corporate reporting tool with a past, present, and future', *The Journal of Business Communication*, Vol. 45, No. 2, pp.94–119. Doi: 10.1177/0021943607313989.

Witten, I., Don, K., Dewsnip, M. et al. (2004) 'Text mining in a digital library', *International Journal on Digital Libraries*, Vol. 4, pp.56–59. Doi: 10.1007/s00799-003-0066-4.

Wu, H.C., Luk, R.W.P., Wong, K,F. and Kwok, K.L. (2008) 'Interpreting TF-IDF term weights as making relevance decisions', *ACM Transactions on Information Systems*, Vol. 26, No. 3. Doi: 10.1145/1361684.1361686.

Zarei, A., Maleki, M., Feiz, D. and Kojuri, M.S. (2018) 'Competitive intelligence text mining: words speak', *Journal of AI and Data Mining*, Vol. 6, No. 1, pp.79-92. Doi: 10.22044/jadm.2017.950.