Emotion Mining and Attribute Selection for Actionable Recommendations to Improve Customer Satisfaction

Angelina A. Tzacheva, Jaishree Ranganathan, Poonam Rajurkar, and Zbigniew W. Ras
Department of Computer Science
University of North Carolina at Charlotte
Charlotte, North Carolina, USA, 28223
Email: aatzache@uncc.edu
Email: jrangan1@uncc.edu
Email: prajurka@uncc.edu
Email: ras@uncc.edu

Abstract—In today's world, business often depends on the customer feedback and reviews. Sentiment analysis helps identify and extract information about the sentiment or emotion of the topic or document. Attribute selection is a challenging problem especially with large datasets in actionable pattern mining algorithms. Action Rule Mining is one of the methods to discover actionable patterns from data. Action Rules are rules that help describe specific actions to be made in the form of conditions that help achieve the desired outcome. The rules help to change from any undesirable or negative state to a more desirable or positive state. In this paper we present a Lexicon based weighted scheme approach to identify emotions from customer feedback data in the area of manufacturing business. Also, we use Rough sets and explore the attribute selection method for large scale datasets. Then we apply Actionable pattern mining to extract possible emotion change recommendations.This kind of recommendations help business analyst to improve their customer service which leads to customer satisfaction and increase sales revenue.

Keywords—Actionable Pattern Discovery, Attribute Selection, Business Data, Data Mining, Emotion.

I. Introduction

According to Merriam-Webster dictionary, emotion is defined as a conscious mental reaction (such as joy, anger or fear) subjectively experienced as strong feeling usually directed toward a specific object and typically accompanied by physiological and behavioral changes in the body [1]. Emotions are integral part of human life. In today's world social media plays and integral role and people communicate their views about a product or business. For example, customer reviews and rating is a major aspect. Positive customer reviews helps business gain more customers. Customer satisfaction determines customer happiness with the products, services and capabilities. Customer satisfaction information, including surveys and ratings, can help a company determine how to best improve their products and services. It is important for the business or organization to understand the voice of customer and get detailed insights of what their customers need and tailor the service accordingly to improve customer loyalty which helps improve business revenue. Emotion mining is the science of identifying human feeling or opinions towards different services, issues and others. Text emotion mining has many applications including helping customer care services, recommending music or movies to users, e-learning materials. In this work we introduce a lexicon based weighted scheme approach for text emotion classification using customer feedback data in the area of manufacturing business. The main goal here is to help business provide better customer service and improve customer satisfaction.

Actionability is a property of the discovered knowledge. Actionable patterns help benefit the user to accomplish the goals. These patterns provide information about past events and can be utilized for prospective decisions. In this work, we discover Actionable Patterns based on the emotions identified from the customer feedback data, which suggest ways to alter the user's emotion to a more positive or desirable state.

Data is distinct pieces of information that contains several features or attributes. Most of the real world data comes with number of attributes attached that are formatted in a specific way and gives meaning to the data. For example - medical data can be high-dimensional with different parameters like symptoms,

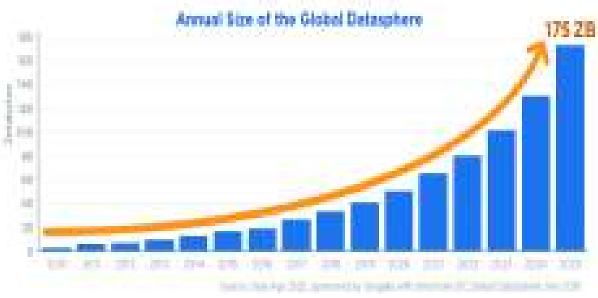

Fig. 1. Annual Size of Global Datasphere [9]

treatment, diagnosis, disease, blood pressure, blood type, medication, admission date, patient information, and many more. The more attributes the data contains, it becomes difficult when dealing with machine learning, data mining, statistical, and or pattern mining algorithms [2], [3].

Thus feature selection has become one of the most important technique in data processing. Some of the attributes in the data may be superfluous or redundant which degenerates the learning algorithm performance. Feature selection is the process of finding a subset of features (that are not redundant or superfluous) from set of all features in the given dataset, while preserving the inherent meaning of the data. Such attributes which fully characterize the knowledge in the database, are called Reducts [4]. This method evaluates the features based on measures such as dependency, distance, information gain, and similarity which are independent of the learning algorithm [5] is called Filter methods, which are considered to be efficient and faster.

Rough set theory [6] is a mathematical approach to deal with lack of certainity or distinctness. Rough set theory provides filter based mathematical frameworks for dimensionality reduction in datasets that uses standard operations in conventional set theory. It uses the existing features in the data and does not require additional parameters to operate. Many methods are for feature selection have been proposed using Rough Sets. Discernibility matrix [7], [8] has been used to generate all possible reducts for the dataset. We are in the era of big data, with astonishing rate of growth of the data. According to International Data Corporation (IDC) [9], worlds data growth rate is at 66% per year which is expected to reach approximately equivalent to 175 zettabytes by the year 2025 (Figure. 1).

To address the BigData challenge, in this work we use Greedy algorithm and Discernibility matrix based algorithm for attribute reduction in distributed cloud frameworks like Spark [10] for scalability and efficiency.

The reminder of the paper is organized as follows: section II - Related work, section III - Background, followed by section IV and V - Methods used and results respectively.

II. Related Work

A. Emotion Mining from Text

Autors Lei et al. [11] present a lexicon-based approach towards social emotion detection. They designed a new algorithm for document selection which has positive effect on the performance of social emotion detection systems. After document selection they exploit words and Part-of-Speech(POS) features. POS helps alleviate the problems of emotional ambiguity of words and the context dependence of the sentiment orientations. Finally generate the emotion lexicon based on the features. They gathered 40,897 news articles assigned with ratings over 8 social emotions including touching, empathy, boredom, anger, amusement, sadness, surprise, and warmness. Their method outperforms the baseline methods of SWAT [12], Emotion-Topic Model (ETM), and Emotion-Term Model (ET) [13] and [14].

Authors Chellal et al. [15] use the categories of informativeness, novelty, and relevance to user interest as conjunctive condition for tweet selection. They propose adaptive method based real time tweet summarization model using the TREC MB RTF-2015 [15] data set.

Authors Aono and Himeno [16] classify tweets with the intensity of emotion based on the idea that n-grams play important role to represent emotions in tweets. They use convolutional neural network model and achieve an average correlation co-efficient of 0.620

Authors Turcu et al. [17] use supervised machine learning models like Naive Bayes, K-Nearest Neighbor and Support Vector Machines, and deep learning neural net tensor flow model, decision tree for affective tweets task and achieve best accuracy of approximately 0.73 for emotion Joy.

B. Attribute Selection and Rough Sets

Author Xie [4], derive boolean matrix directly from the information system and generate set of reducts. The core idea of this algorithm is to use appearance regularity of the elements of reductions and supersets of reductions to construct the set of all reducts. Similarly in [18], initial information system or table is transformed into a special decision table in the form of matrix. Then by using set of simplification rules on the derived table they construct the set of reducts using dynamic programming.

It is applied to medium small sized data tables from UCI Machine Learning repository [19].

Another traditional approach [20] is using attribute consistency check for finding the reduct and core of consistent datasets. This algorithm removes one condition attribute at a time to check for consistency of the remaining table. A table is said to be consistent if for two or more rows or cases for the same values of condition attributes we have the same decision otherwise it is inconsistent. Thus the attributes that are required for the consistent table are marked as reducts of the dataset. This approach is extensive and it is only applicable for smaller datasets.

Authors Wang et al. [21] and Author Skowron [8] use discernibility matrix based algorithm for finding attribute reducts. Similarly [7] propose a discernibility based funtion to find reducts. Authors Al-Radaideh et al., [22], also use the discernibility matrix modulo as input for the heuristic reduct computation approach by attributes weighting. They use different attribute weights such as global weight, local weight, and attribute value cardinalities for the purpose of reduct generating algorithm.

Authors Korzen and Jaroszewicz [23] find reducts without explicitly building the discernibility matrix, by use of conditional Gini index. The discernibility based algorithms are usually expensive especially for large datasets. There have been many heuristic and greedy algorithms proposed for attribute reduction.

Rough set theory was first introduced by Zdzislaw Pawlak [24]. This approach of Rough Sets constitutes a base for Knowledge Discovery in Databases, Machine Learning, Decision Support Systems, Pattern Recognition. Some of the problems approached using rough set theory are feature selection, eliminate redundant data, identifiy data dependencies, pattern extraction like association rules, discover data similarities or differences, and approximate data classification. Rough set based methods have been applied in the area of business, web and text mining [25], image processing [26], medicine, bioinformatics [27], economics [28].

C. Actionable Pattern Mining

Data may have interesting patterns and insights which can be discovered using Action Rule Mining method. Action rules are rules that provide suggestions or actions on possible state change that benefits the user.

Actionable patterns or Action rules play an important role in the applications of Emotion mining. Some examples include: suggest calming music or mood enhancing movies in smart phones, car, and other such systems where the application identifies the user emotion and suggest ways to enhance negative emotion to more positive emotion. Authors in [29], [30] identify emotions in tweet data and extract action rules for emotion change recommendation.

III. Background

In this section we describe basic terms including information system, decision table, and discernibility matrix associated with Rough sets.

A. Information System

An information System is denoted by $\mathbb{I} = \langle \mathbb{U}, \mathbb{A}, \mathbb{V} \rangle$, where $U = \{x_1, x_2, ..., x_n\}$ is finite, nonempty set of objects; $A = \{a_1, a_2, ..., a_m\}$ is finite, nonempty set of attributes; V is the domain of attribute set A. Let us consider a sample information system in Table. I, where $\mathbb{U}$ is set of objects $\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8\}$ in the system; $\mathbb{A}$ is non-empty finite set of attributes $\{A, B, C, E, F, G, D\}$; $\mathbb{V}$ is the domain of attributes in $\mathbb{A}$, for instance the domain of attribute B in the system $\mathbb{Z}$ is $\{B_1, B_2, B_3\}$.

TABLE I
INFORMATION SYSTEM I

X	A	B	C	E	F	G	D
x_1	A_1	B_1	C_1	E_1	F_2	G_1	D_1
x_2	A_2	B_1	C_2	E_2	F_2	G_2	D_3
x_3	A_3	B_1	C_1	E_2	F_2	G_3	D_2
x_4	A_1	B_1	C_2	E_2	F_2	G_1	D_2
x_5	A_1	B_2	C_1	E_3	F_2	G_1	D_2
x_6	A_2	B_1	C_1	E_2	F_3	G_1	D_2
x_7	A_2	B_3	C_2	E_2	F_2	G_2	D_2
x_8	A_2	B_1	C_1	E_3	F_2	G_3	D_2

B. Decision Table

The information system $\mathbb{I}$ is called a decision table $\mathbb{DT}$, when it contains a distinguished decision attribute D and condition attributes C. Decision table is denoted by $\mathbb{DT} = \langle \mathbb{U}, \mathbb{A}, \mathbb{V} \rangle$, where U and V are same as defined in section III-A, and A = $C \cup D$. The information system in Table.I, is decision system if the attributes $\mathbb{A}$ are classified into Conditional C, and Decision D, $\mathbb{A} = (C \cup D)$, where C = $\{A, B, C, E, F, G\}$, and D = D.

C. Discernibility Matrix

Let us consider the information system described in section III-A, $\mathbb{DM}(\mathbb{IS})$ - discernibility matrix of the information system $\mathbb{I}$ is represented by a $n \times n$ matrix c_{ij}, given by Eq. 1.

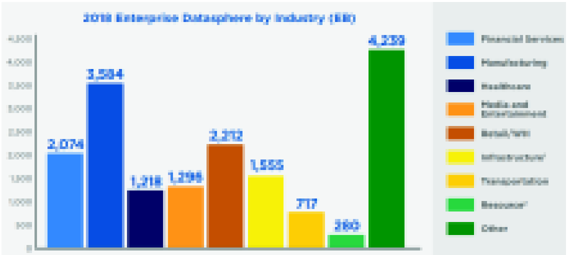

Fig. 2. 2018 Enterprise Datasphere by Industry [9]

$$c_{ij} = \{a \in A : a(x_i) \neq a(x_j)\}, for i, j = 1, 2, ..., n. \tag{1}$$

Similarly the discernibility matrix $\mathbb{DM}(\mathbb{DT})$, for a decision table (section III-B) is denoted by a $n \times n$ matrix c_{ij}, given by Eq. 2.

$$c_{ij} = \begin{cases} \phi & f_d\ (x_i) = f_d\ (x_j); \\ \{a \in A : a(x_i) \neq a(x_j)\} & f_d\ (x_i) \neq f_d\ (x_j) \end{cases} \tag{2}$$

The discernibility matrix $\mathbb{DM}(\mathbb{DT})$, for a decision table (section III-B) is given in Table. II.

TABLE II
DISCERNIBILITY MATRIX DT

X	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1								
x_2	ACEG							
x_3	AEG	ACG						
x_4	CE	ACG	ϕ					
x_5	BE	ABCEG	ϕ	ϕ				
x_6	AEF	CFG	ϕ	ϕ	ϕ			
x_7	ABCEG	B	ϕ	ϕ	ϕ	ϕ		
x_8	AEG	CEG	ϕ	ϕ	ϕ	ϕ	ϕ	

D. Indiscernibility Relation - Equivalence Relation

Let us consider the Information System $\mathbb{I} = \langle \mathbb{U}, \mathbb{A}, \mathbb{V} \rangle$. For every set of attributes $B \subseteq A$, an equivalence relation is denoted by $IND_A(B)$ and called the B-indiscernibility relation, which is given in Eq. 3.

$$IND_A(B) = \{(u, u') \in U^2 : a \in B, a(u) = a(u')\} \tag{3}$$

IV. METHODOLOGY

In this section we present our proposed approach of Emotion mining and attribute reducts for actionable pattern mining.

A. Data Collection

According to the IDC [9], 48% of enterprise data sphere comprises of data from the Manufacturing, Health-care, Financial services, Media and Entertainment industries. Out of the four industries, it is notable from Fig. 2 that Manufacturing is responsible for largest share of data.

Given the amount of data in the manufacturing industry, in this study we use the Net Promotor Score (NPS) Data [31] collected by telephone surveys on customer satisfaction. Net Promotor Score is a standard metric used for measuring customer satisfaction by labeling customers as ‘Promotor’, ‘Passive’, or ‘Detractor’.The actual dataset was collected in a span of years. In this study we use the Text feedback from customers collected during the years 2015 - 2016 along with other features containing customer details, survey details, and benchmark questions on which the service is being evaluated. The surveys are for 38 companies located across United States and parts of Canada. We process the text feedback and derive the decision attribute as Customer Emotions. The decision problem is to improve customer satisfaction/loyalty by identifying factors to improve customer emotions. If a customer feels joy, or trust then it is highly likely that the customer remains loyal to the business and also improves Net Promotor Score.

B. Data Processing and Feature Augmentation

The data collected consists of text comments in the form of customer survey in the business data. The text data is pre-processed as in Fig. 3

Text feedback is not syntactically well formed. So the text need considerable pre-processing before further steps of emotion labeling, and feature extraction. We lowercase all the letters in the text, remove stop words, replace slang words with formal text.

We add a new feature called emotion to the data. As part of emotion labeling we use the National Research Council - NRC Emotion lexicon [32]–[35].

The Annotations in the lexicon are at WORD-SENSE level. Each line has the format: <Term> <AffectCategory> <AssociationFlag>. *Term* indicates the word for which the emotion associations are provided, *AffectCategory* indicates one of the eight emotions (‘anger’, ‘fear’, ‘disgust’, ‘sadness’, ‘anticipation’, ‘joy’, and ‘trust’), and *AssociationFlag* indicates one of two possible associations either ‘0’ (target word has no association with affect category) or ‘1’ (target word has association with affect category).

TABLE III
DATASET PROPERTIES: BUSINESS DATA - NET PROMOTOR SCORE (NPS)

Property	**NPS Business Data: Client Comments Parts 2015**	**NPS Business Data: Client Comments Parts 2016**	**NPS Business Data: Client Comments Service 2015**	**NPS Business Data: Client Comments Service 2016**
Attributes	23 attributes including - Client Name - Division - SurveyType - ChannelType - BenchmarkAll: DealerCommunication -BenchmarkAll: LikelihoodtobeRepeatCustomer - Emotion	37 attributes including - Client Name - Division - SurveyType - ChannelType - BenchmarkAll: ContactStatusofFutureNeeds - BenchmarkParts: AvailabilityYN - BenchmarkParts: EaseofCompletingPartsOrder - Emotion	24 attributes including - Client Name - Division - SurveyType - ChannelType - Benchmark: All - Ease of Contact - Benchmark: Service - Repair Completed Correctly - Benchmark: Service - Repair Completed Timely - Emotion	38 attributes including - Client Name - Division - SurveyType - ChannelType - Benchmark: All - Contact Status of Future Needs - Benchmark: All - Contact Status of Issue - Benchmark: Dealer Offers solutions to support the success of cus - Benchmark: Referral Behavior - Emotion
# of instances	6656	10102	11121	17706

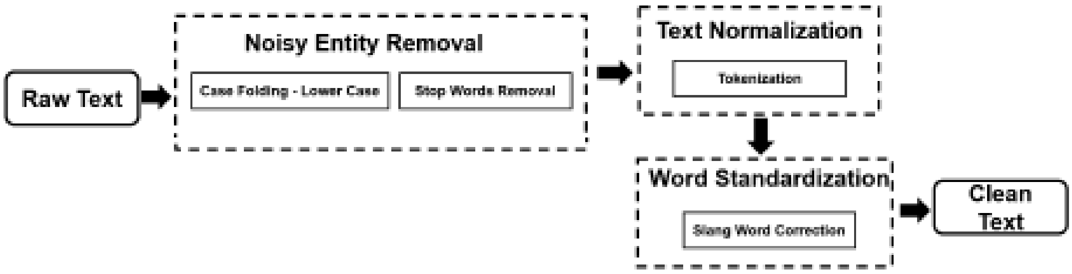

Fig. 3. Text Pre-Processing

Each text feedback in the dataset is annotated with the corresponding emotion based on the weights computed using features such as word-sense. Thus we obtain annotated dataset where each record is labeled with corresponding emotions such as 'anger', 'fear', 'disgust', 'sadness', 'anticipation', 'joy', and 'trust'.

C. Rough Sets for Attribute Reduction

1) Computation of Reducts using Discernibility Matrix: This method computes the discernibility matrix [36] for the given dataset, which is then used for computing reducts [37].

2) Computation of Reducts using Quick Reduct Algorithm: We use the QuickReduct [38] algorithm, here the computation time depends on the number of attributes instead of the number of objects. Thus it runs faster than the one computation method. which uses the Discernibility Matrix section IV-C1.

Algorithm 1: Reduct Computation

Input: Information System A
Output: $Reduct_A$(A)
compute the indiscernibility matrix M(A) = (C_{ij})
Reduce M
d - number of non-empty fields of reduced M
Build set of Reducts R_0**,** R_1**, ... ,** R_d
Remove the redundant elements from R_d
RED_A**(A) =** R_d

D. SparkR

SparkR [39] is an R package for the use of Apache Spark with R. Spark 2.4.4, SparkR provides a distributed data frame implementation that supports operations like selection, filtering, aggregation etc. (similar to R data frames, dplyr) but on large datasets. SparkR also sup-

Algorithm 2: Quick Reduct Algorithm

Input: Decision Table A
Output: Superreduct R
R $\leftarrow$ **{} ;**
repeat the following
T $\leftarrow$ **R ;**
foreach x $\in$ **(A - R) do**
if $\gamma_{R \cup \{x\}} > \gamma_T$ **then T** $\leftarrow$ **R** $\cup$ **{x};**
R $\leftarrow$ **T;**
end
until γ_R **==** γ_A

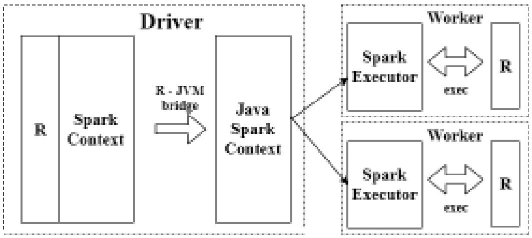

Fig. 4. SparkR Architecture [39]

ports distributed machine learning using MLlib. The SparkR architecture Fig. 4, consists of R to JVM brige on the driver for submitting jobs to cluster and spark executor on the worker for running R. For the purpose of our experiements we use the SparkR Roughsets package [40].

E. Actionable Pattern Mining

Action Rules mining is a method to extract actionable patterns from the data. These rules helps identify possible state change (from one state to another), to a more desirable state. Consider the Information System in Table.I. This Information System is denoted as Decision system if the attributes $\mathbb{M}$ are classified into flexible M_{fl} , stable M_{st} and decision d, $\mathbb{M} = (M_{st}, M_{fl}, \{d\})$. From Table. I $M_{st} = \{A, B, C\}$, $M_{fl} = \{E, F, G\}$, and $d = D$.

The expression $(\mathrm{y}, \mathrm{y}_1 \rightarrow \mathrm{y}_2)$ is an atomic action term, where y is an attribute and y_1, $\mathrm{y}_2 \in \mathrm{V}_y$. If $\mathrm{y}_1 = \mathrm{y}_2$), then y is stable on y_1. In this case action term is denoted as $(\mathrm{y}, \mathrm{y}_1)$ for simplicity.

- If t is an atomic action term, then t is an action term.
- If t_1 , t_2 are action terms, then $\mathrm{t}_1 * \mathrm{t}_2$ is an action term.
- If t is an action term containing $(\mathrm{y}, \mathrm{y}_1 \rightarrow \mathrm{y}_2)$, $(\mathrm{z}, \mathrm{z}_1 \rightarrow \mathrm{z}_2)$ as its sub-terms, then $\mathrm{y} \neq \mathrm{z}$.
- Domain of action term is denoted by $\mathrm{Dom(t)}$, which includes all attributes listed in t.

The expression $\mathrm{r} = [\mathrm{t}_1 \rightarrow \mathrm{t}_2]$ is an Action Rule where, t_1 is an action term and t_2 is an atomic action term. Eq. 4 is an example Action Rule from Table.I.

$$[B_1 \wedge C_1 \wedge (F, F_3 \rightarrow F_1) \wedge (G, \rightarrow G_1) \longrightarrow (D, D_2 \rightarrow D_1)] \tag{4}$$

Support and confidence of rule r is given as below:

- $sup(r) = min\{card(Y_1 \cap Z_1), card(Y_2 \cap Z_2)\}$.
- $conf(r) = \frac{card(Y_1 \cap Z_1)}{card(Y_1)} \cdot \frac{card(Y_2 \cap Z_2)}{card(Y_2)}$.
- $card(Y_1) \neq 0$, $card(Y_2) \neq 0$, $card(Y_1 \cap Z_1) \neq 0$, $card(Y_2 \cap Z_2) \neq 0$.
- $conf(r) = 0$ otherwise.

V. Experiments and Results

In this section we use the original raw data explained in section IV-A apply the following methods: Emotion Labeling (section. IV-B) on the customer feedback - text data, that labels the text into eight basic emotions ('joy', 'sadness', 'surprise', 'trust', 'anticipation', 'disgust', 'fear', 'anger'); attribute reduct (section. IV-C); and Actionable Pattern Discovery.

A. Emotion Labeling

After Emotion Labeling, the Net Promotor Score Business Data contains the following features as mentioned in Table. IV.

TABLE IV
Business Data - Features.

Dataset	# Instances	# Attributes
Client Comments Parts 2015	6656	23
Client Comments Service 2015	11121	22
Client Comments Parts 2016	10102	37
Client Comments Service 2016	17706	36

B. Reducts using Rough Sets

The experiments are performed in both Single node local machine and University Research Cluster (URC), which includes the following services: HBase, Hive, Hue, Impala, Kudu, Oozie, Spark, Spark2, Sqoop2, and YARN. It has 16 nodes with dual Intel 2.93 GHz 6-core processors. The data in Table. IV is used to run two reduct computation algorithms using rough sets explained in section IV-C1 and section IV-C2.

TABLE V
BUSINESS DATA - QUICK REDUCT.

Dataset	Single Node - approx. Time Taken	URC Cluster - approx. Time Taken
Client Comments Parts 2015	36.3 seconds	32.74 seconds
Client Comments Service 2015	1.05 mins	1.102 mins
Client Comments Parts 2016	3.116 mins	3.092 mins
Client Comments Service 2016	6.557 mins	6.26 mins

TABLE VI
BUSINESS DATA - DISCERNIBILITY MATRIX - SINGLE NODE.

Dataset	Single Node approx. Time Taken (mins)	
	Discernibility Matrix	Reduct Computation
Client Comments Parts 2015	4.18	4.63
Client Comments Service 2015	50.85	8.93
Client Comments Parts 2016	23.18	12.70
Client Comments Service 2016	72	157.8

TABLE VII
BUSINESS DATA - DISCERNIBILITY MATRIX - URC CLUSTER.

Dataset	URC Cluster approx. Time Taken (mins)	
	Discernibility Matrix	Reduct Computation
Client Comments Parts 2015	2.99	2.28
Client Comments Service 2015	59.05	8.75
Client Comments Parts 2016	7.6	981secs
Client Comments Service 2016	184.8	70.8

C. Action Rule

We generate Action Rules for the Business Data labeled with emotions, generated in section V-A. The dataset consists of records describing attributes of Machinery parts company sales, including the customer feedback text comments and their corresponding Emotion in one of the following category ‘joy’, ‘trust’, ‘anticipation’, ‘fear’, ‘disgust’, ‘sadness’, ‘anger’, ‘surprise’. We choose Emotion as the decision attribute and generate Action Rules that help identify changes that are required for the Emotion to be more positive. For example, to change the emotion from ‘anticipation’ to ‘trust’.

The decision problem here is to suggest possible recommendations to the Machinery parts companies sales, on how to make customers feel much better which ultimately helps improve customer loyalty and satisfaction.

The data is first processed to contain only attributes from the attribute reduction step. Later the data is divided into two parts with decision attribute ‘Emotion’ on each part of the data for distributed processing. Then combined to obtain the final set of rules. Extracted sample Action Rules are provided in Table. VIII.

Let us consider the rule ClientCommentParts2015$_{\mathrm{AR2}}$ in Table. VIII. According to the Action Rule, if the Benchmark value of time it took to place the order ($BenchmarkPartsTimeitTooktoPlaceOrder$) is changed from 9 to 10 and Overall Satisfaction and Order Accuracy values are maintained at 10, then it is possible to gain customer Trust.

VI. CONCLUSIONS AND FUTURE WORKS

This paper presents application of identifying Emotions from the customer survey feedback in the manufacturing Industry data and applying Action Rule mining methods to identify actionable patterns to help improve customer satisfaction. Most of the work to build intelligent data mining systems, machine learning algorithms, pattern mining tend to face the bottleneck at the point of the acquiring best possible knowledge to extract useful patterns. In this paper we use the rough-set attribute reduction algorithms, concept of SparkR distributed framework to overcome the problem of high dimension data and chose the best possible attributes for pattern discovery. We apply our method on domains, especially that are generating high volume of data in the current world: Business data. The execution time on Spark cluster is much less compared to the Single Node machine. In future we plan to improve attribute selection process by splitting the data into sub-systems, using intelligent data partitioning and attribute weightage schemes.

REFERENCES

[1] M. W. Staff, “Online dictionary: http://www. mw. com/home. htm, 2004. merriam–webster, inc., 47 federal street,” *PO Box*, vol. 281.

[2] B. Raman and T. R. Ioerger, “Instance-based filter for feature selection,” *Journal of Machine Learning Research*, vol. 1, no. 3, pp. 1–23, 2002.

[3] L. C. Molina, L. Belanche, and À. Nebot, “Feature selection algorithms: A survey and experimental evaluation,” in *2002 IEEE International Conference on Data Mining, 2002. Proceedings.* IEEE, 2002, pp. 306–313.

[4] Y. Xie, “An attribute reduction method based on boolean matrix,” in *2008 International Conference on Computer Science and Software Engineering*, vol. 1. IEEE, 2008, pp. 423–426.

[5] W. Shu and H. Shen, “Incremental feature selection based on rough set in dynamic incomplete data,” *Pattern Recognition*, vol. 47, no. 12, pp. 3890–3906, 2014.

[6] Z. Pawlak, “Rough set theory and its applications to data analysis,” *Cybernetics & Systems*, vol. 29, no. 7, pp. 661–688, 1998.

[7] S. K. Pal and A. Skowron, *Rough-fuzzy hybridization: A new trend in decision making*. Springer-Verlag, 1999.

[8] A. Skowron, “Extracting laws from decision tables: a rough set approach,” *Computational Intelligence*, vol. 11, no. 2, pp. 371–388, 1995.

TABLE VIII
SAMPLE ACTION RULES - BUSINESS DATA.

Enhance Customer Emotion - Anticipation → Trust
1) $ClientCommentParts2015_{AR1}$: $(BenchmarkAllOverallSatisfaction = 10) \wedge (BenchmarkPartsExplanationofDeliveryOptionsCosts = 10) \wedge (BenchmarkPartsHowOrdersArePlaced, 2 \rightarrow 3) \Longrightarrow (Emotion, Anticipation \rightarrow Trust)[Support : 10, Confidence : 9.21\%]$
2) $ClientCommentParts2015_{AR2}$: $(BenchmarkAllOverallSatisfaction = 10) \wedge (BenchmarkPartsOrderAccuracy = 10) \wedge (BenchmarkPartsTimeitTooktoPlaceOrder, 9 \rightarrow 10) \Longrightarrow (Emotion, Anticipation \rightarrow Trust)[Support : 7, Confidence : 8.88\%]$
3) $ClientCommentParts2016_{AR1}$: $(BenchmarkAllOverallSatisfaction = 10) \wedge (BenchmarkPartsPartsAvailability = 10) \wedge (BenchmarkPartsPromptNotificationofBackOrders = 10) \Longrightarrow (Emotion, Anticipation \rightarrow Trust)[Support : 151, Confidence : 13.81\%]$

[9] D. Reinsel, J. Gantz, and J. Rydning, "Data age 2025: the digitization of the world from edge to core," *Seagate https://www. seagate. com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper. pdf*, 2018.

[10] M. Zaharia, M. Chowdhury, M. J. Franklin, S. Shenker, and I. Stoica, "Spark: Cluster computing with working sets." *HotCloud*, vol. 10, no. 10-10, p. 95, 2010.

[11] J. Lei, Y. Rao, Q. Li, X. Quan, and L. Wenyin, "Towards building a social emotion detection system for online news," *Future Generation Computer Systems*, vol. 37, pp. 438–448, 2014.

[12] P. Katz, M. Singleton, and R. Wicentowski, "Swat-mp: the semeval-2007 systems for task 5 and task 14," in *Proceedings of the 4th international workshop on semantic evaluations*. Association for Computational Linguistics, 2007, pp. 308–313.

[13] S. Bao, S. Xu, L. Zhang, R. Yan, Z. Su, D. Han, and Y. Yu, "Joint emotion-topic modeling for social affective text mining," in *Ninth IEEE International Conference on Data Mining, ICDM'09.* IEEE, 2009, pp. 699–704.

[14] S. Bao, S. Xu, L. Zhang, R. Yan, Z. Su, D. Han, and Y. Yu, "Mining social emotions from affective text," *IEEE transactions on knowledge and data engineering*, vol. 24, no. 9, pp. 1658–1670, 2012.

[15] A. Chellal, M. Boughanem, and B. Dousset, "Multi-criterion real time tweet summarization based upon adaptive threshold," in *Proceedings of 2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI).* IEEE, 2016, pp. 264–271.

[16] M. Aono and S. Himeno, "Kde-affect at semeval-2018 task 1: Estimation of affects in tweet by using convolutional neural network for n-gram," in *Proceedings of The 12th International Workshop on Semantic Evaluation*, 2018, pp. 156–161.

[17] R.-A. Turcu, S. M. Amarandei, I.-A. Flescan-Lovin-Arseni, D. Gifu, and D. Trandabat, "Emointens tracker at semeval-2018 task 1: Emotional intensity levels in# tweets," in *Proceedings of The 12th International Workshop on Semantic Evaluation*, 2018, pp. 177–180.

[18] H. AbouEisha, "Finding optimal exact reducts," in *Proceedings of the International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management-Volume 1.* SCITEPRESS-Science and Technology Publications, Lda, 2014, pp. 149–153.

[19] D. Dua and C. Graff, "UCI machine learning repository," 2017. [Online]. Available: http://archive.ics.uci.edu/ml

[20] R. Vashist, "An algorithm for finding the reduct and core of the consistent dataset," in *2015 International Conference on Computational Intelligence and Communication Networks (CICN).* IEEE, 2015, pp. 741–745.

[21] R. Wang, D. Miao, and G. Hu, "Discernibility matrix based algorithm for reduction of attributes," in *Proceedings of the 2006 IEEE/WIC/ACM international conference on Web Intelligence and Intelligent Agent Technology*. IEEE Computer Society, 2006, pp. 477–480.

[22] Q. A. Al-Radaideh, M. N. Sulaiman, M. H. Selamat, and H. Ibrahim, "A heuristic reduct computation approach by attributes weighting for rough set based classification," *Journal of Computer Sciences (Special Issue)*, vol. 41, p. 47, 2005.

[23] M. Korzen and S. Jaroszewicz, "Finding reducts without building the discernibility matrix," in *5th International Conference on Intelligent Systems Design and Applications (ISDA'05).* IEEE, 2005, pp. 450–455.

[24] Z. Pawlak, "Rough sets," *International journal of computer & information sciences*, vol. 11, no. 5, pp. 341–356, 1982.

[25] J. Watada, S. C. Tan, Y. Matsumoto, and P. Vasant, "Rough set-based text mining from a large data repository of experts diagnoses for power systems," in *International Conference on Intelligent Decision Technologies*. Springer, 2017, pp. 136–144.

[26] G. Jothi *et al.*, "Hybrid tolerance rough set–firefly based supervised feature selection for mri brain tumor image classification," *Applied Soft Computing*, vol. 46, pp. 639–651, 2016.

[27] L. Tang and J. Wang, "Predicting enzyme class with rough sets," in *2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI).* IEEE, 2016, pp. 1832–1836.

[28] F. E. Tay and L. Shen, "Economic and financial prediction using rough sets model," *European Journal of Operational Research*, vol. 141, no. 3, pp. 641–659, 2002.

[29] J. Ranganathan, A. S. Irudayaraj, and A. A. Tzacheva, "Action rules for sentiment analysis on twitter data using spark," in *Data Mining Workshops (ICDMW), 2017 IEEE International Conference on*. IEEE, 2017, pp. 51–60.

[30] J. Ranganathan and A. Tzacheva, "Emotion mining in social media data," *Procedia Computer Science*, vol. 159, pp. 58 – 66, 2019, knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 23rd International Conference KES2019.

[31] Z. W. Ras, K. A. Tarnowska, J. Kuang, L. Daniel, and D. Fowler, "User friendly nps-based recommender system for driving business revenue," in *Proceedings of 2017 International Joint Conference on Rough Sets (IJCRS'17), LNCS*, vol. 10313. Springer, 2017, pp. 34–48.

[32] S. M. Mohammad and P. D. Turney, "Crowdsourcing a word–emotion association lexicon," *Computational Intelligence*, vol. 29, no. 3, pp. 436–465, 2013.

[33] S. M. Mohammad and P. D. Turney, "Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon," in *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text*. Association for Computational Linguistics, 2010, pp. 26–34.

[34] S. M. Mohammad and S. Kiritchenko, "Using hashtags to capture

fine emotion categories from tweets," *Computational Intelligence*, vol. 31, no. 2, pp. 301–326, 2015.

[35] S. M. Mohammad, "# emotional tweets," in *Proceedings of the First Joint Conference on Lexical and Computational Semantics-Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation*. Association for Computational Linguistics, 2012, pp. 246–255.

[36] A. Skowron and C. Rauszer, "The discernibility matrices and functions in information systems," in *Intelligent decision support*. Springer, 1992, pp. 331–362.

[37] J. G. Bazan, H. S. Nguyen, S. H. Nguyen, P. Synak, and J. Wróblewski, "Rough set algorithms in classification problem," in *Rough set methods and applications*. Springer, 2000, pp. 49–88.

[38] Q. Shen and A. Chouchoulas, "A modular approach to generating fuzzy rules with reduced attributes for the monitoring of complex systems," *Engineering Applications of Artificial Intelligence*, vol. 13, no. 3, pp. 263–278, 2000.

[39] S. Venkataraman, Z. Yang, D. Liu, E. Liang, H. Falaki, X. Meng, R. Xin, A. Ghodsi, M. Franklin, I. Stoica *et al.*, "Sparkr: Scaling r programs with spark," in *Proceedings of the 2016 International Conference on Management of Data*. ACM, 2016, pp. 1099–1104.

[40] L. S. Riza, A. Janusz, C. Bergmeir, C. Cornelis, F. Herrera, D. Śle, J. M. Benítez *et al.*, "Implementing algorithms of rough set theory and fuzzy rough set theory in the r package "roughsets"," *Information Sciences*, vol. 287, pp. 68–89, 2014.