JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Assessing the Quality of Scientific Documents by Using Machine Learning Techniques

Aditi Bhardwaj*[1], Vikash Singh Rajput[1]

Department of computer science, ShriRam College of Engineering & Management, Banmore, India

1. Name: Prof. Vikash Singh Rajput

Address: Department of computer science, ShriRam College of Engineering & Management, Banmore, Morena, (M.P.)-476444, India

2. Name: Ms. Aditi Bhardwaj (PG Scholar)

Address: Department of computer science, ShriRam College of Engineering & Management, Banmore, Morena, (M.P.)-476444, India

Abstract

This was accomplished by employing an established finding method that involves the steps of data collection, preliminary processing, selection of a suitable data mining technique, and interpretation of the trends discovered. The findings were then applied to more study. The sentence-level breakdown of the written material database resulted in nine sets of data, each of which has 22,327 data items for the two different interpretation types (automatic interpretation and expert translation). One or two automated translation methods were utilized for translating the source text and provide references for the provided candidate texts in order to construct a standard translation. Five artificial intelligence algorithms-Decision Trees, k-Nearest Neighbor, Artificial Neural Networks, Naive Bayes, and SVM were applied in several iterations with the preprocessed dataset. After learning the algorithms, the methods' variables were tuned and they were evaluated on an exception set that was extracted from the database. The k-Nearest Neighbor classification produced the best outcomes, scoring 73 when given knowledge to the original file and 63% when not. The optimised methods were applied to the words of each initial record to make a claim about paragraph level categories, and the outcomes were then re-combined to categorize the relevant documents.

Keywords: Machine learning, Data mining, Decision tree, Support vector machine, Convolutional neural network, Translation

Introduction

Since the start of the sixteenth century, overcoming linguistic obstacles has been a hot issue. From then, tools and concepts, including robotic dictionaries or global languages, have emerged to facilitate communication between speakers of various tongues [1-3]. The capability of automatically converting documents from a single tongue to a different one without any human involvement is an issue that have been studied for around 60 years and has grown in importance during the past ten years because of increasingly internationalized businesses and general globalization. Businesses must offer professionally translated product information if they want to succeed in a global market. There may be several separate papers for one item due to the fact that sophisticated products sometimes have multiple user groups, such as managers, consumers, and programmers [4-6]. Finding qualified translators with the necessary knowledge of technology to provide accurate translations of technical material at fair prices is challenging, particularly for businesses that focus on exporting. So businesses are becoming more interested in automated translation services.

It is crucial to offer text in computerized, high-quality versions in order to guarantee equitable information availability no matter the tongue of the original content. Businesses place a great importance on correct technological translations because they are necessary for efficient processes, client fulfillment, and they depict the care, functioning, and security aspects of goods. Inconsistencies in this area can have serious consequences. Additionally, miscommunications brought on by poor translations might hurt business partnerships [7].

The output produced by software for machine translation continues to be validated in order to assure the needed quality because machine learning for language processing is a highly complicated topic [8-9]. Thus, using automation for interpreting business papers only shifts the issue to the document's generation to its examination and repair but doesn't solve it. As a result, evaluating translated technical information is a crucial step for businesses to cut both expenses and time and develop an efficient method of translating vital material. This also guarantees a particular standard of quality. Because to the personal nature of the word "quality" and its various connotations, such as grammatical accuracy, stylistic refinements, or logical accuracy, judging the quality of translation can be challenging.

It is yet futuristic to have possession of electronic systems that can accurately translate every statement, particularly given the difficulty of conveying the meaning of a phrase to the machine. Regarding this issue, the ability to rate the standard of a particular translations is crucial since without it, it would be impossible to guarantee that an article has been accurately rendered [10]. The huge volume of scientific documentation seen in any business that sells items makes the emphasis on this form of documentation particularly intriguing and serves as an incentive to motivate automatic translation of this sort of material. These days, businesses outsource their technological documentation translation needs to outside translators in order to solve the issue. Since the individual ordering the translations may not be a native speaker of the target language, it is crucial to confirm whether the job was completed correctly and expertly by an individual and not by an automatic translating system [11-13].

Objectives and research issues

In this thesis, a machine learning technique will be used in a knowledge discovery process to classify documents by their translation type (professional translation, automated translation). Further, an approach on how to evaluate the quality of translated technical documents will be proposed. Concerning this issue, we address two main research questions: 1) How can the translation quality of technical documents be evaluated, given the original document is available? 2) How can the translation quality of technical documents be evaluated, given the original document is not available?

Methods

Data mining

DKD is frequently interchanged with data mining. This is really the step in the DKD process when the right strategy and algorithms are picked and then used for the data collection [10]. As a result, it plays a crucial role in the procedure of data retrieval in databases [13]. Data mining is the process that involves applying analytical techniques to any type of data with the goal to uncover trends or structures in the information collection and then utilizing such structures to categorize the information into multiple groups (labels). Computer structures, data, and recognition of patterns are some of the study areas that it includes.

Machine learning (ML)

Various researches have long used machine learning. Convolutional neural networks (CNN) and decision tree structures (DTs) have been successfully used for approximately 20 years in the detection and monitoring of many data sets [15-18]. MLAs are used now for a variety of tasks, from identifying and characterizing datasets using methods to defining originality using [20–22]. In fact, the majority of the PubMed data and almost 1500 papers on documents translation utilizing machine learning have been published. These publications' use of machine learning approaches to the detection, characterization, tracking, or differentiation of datasets and other kinds of data, however, constitutes the great bulk of their topics. Alternately, machine learning is primarily utilized in the identification and prevention of data duplicity [23]. Recently, experts in data mining have tried to employ machine learning to identify and predict the documentation. This section covers the widely used machine learning algorithms drawn from the aforementioned types of machine learning.

Bayesian Network (BNs)

A subject that cannot be identified can be expressed and thought about using Bayesian Networks. The theory of probability and recognition of patterns are combined in BNs. BNs represent the probabilistic model controlling a number of indicators by itemizing a collection of dependent assumptions of independence as well as a group of likelihood functions. The network's network is made up of either continuous or discrete parameterss, and arcs show how they are related to one another. Conditional independence is a key idea in Bayesian Networks [29].

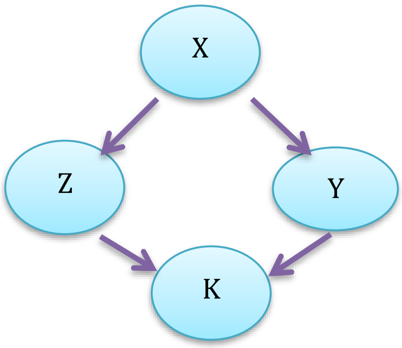

Figure 1 The framework comprises of a Bayesian Network

The probability of K is independent and equally distributed on X given the likelihoods of Z, Y, and the arrangement in Figure 1, as indicated in the following equation:

$$P(K \mid X, Z, Y) = P(K \mid Z, Y)$$

If a vertex's quick antecedents are present, it is said to be absolutely independent of quasi. As a result, it is easier to comprehend the intended attribute.

K-Nearest Neighbors (KNNs)

It is one of the best algorithms and methods for problems with regression and classification. KNN algorithms primarily use similarity tests to distinguish new information from locations in news streams. A majority vote among the area's neighbors determines its classification [30].

Naive Bayes (NBs)

It focuses on probabilistic methods that compare one characteristic of a category with another that has various beliefs and characteristics that could be quite different from one another and might. It is believed that the category with the best likelihood of success will be the most ideal. Nave Bays might be thought of as a specific instance of a Bayesian network re 3) [31].

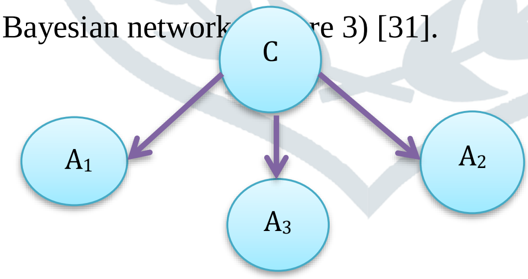

Figure 2 Nave Bays a unique situation

If an element in figure 2 (A1) and another element (C1) are in the same category (C), then they are both independent events. The Bayesian Network's global supposition limits the parameters by defining a set of conditional independence assumptions between the features and a collection of likelihood functions. The Naive Bayes presupposition, on either extreme, appears to be more limiting than this.

Random Forest (RF)

I It is employed in the creation of many different decision trees. Based on the quantity of leaves, Random Forest selects the final outcome [32].

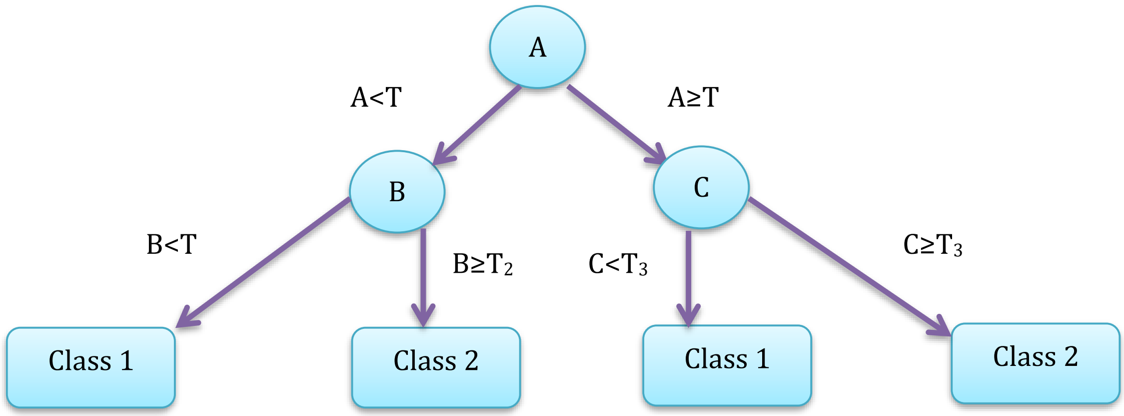

Figure 3 Algorithms for Random Forests

The tree-like structure of DTs is seen in Figure 3. Every variable (A, B, and C) is represented by a circle, while the selection outcomes (Class 1, Class 2) are represented by squares. T(1-3) denotes the thresholds (classification criteria) required to assign every parameters to a class mark.

Convolutional Neural Network (CNNs)

Convolutional neural networks (CNNs) are a mode of reasoning inspired by the human brain [33]. With more academics employing them, ANNs have emerged as a popular study area in recent years. A number of milestones were also produced by CNNs, including important developments in the early phases of BC diagnosis and categorization [34]. Input, secret, and output are the three layers that typically make up a CNNs model (Figure 4) [35]. Despite being employed, the BP method has a number of drawbacks when working with big volumes of data. A revised BP method is rarely used in operational utilization because BP calculations are time-consuming and need extensive planning.

A

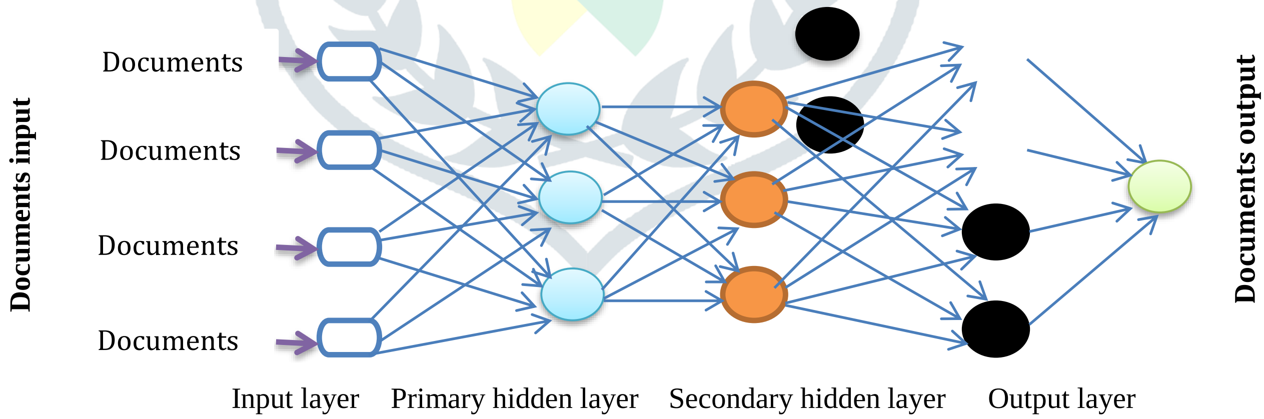

Figure 4 A brief overview of the training of a CNN to predict diagnostic results using four inputs, two hidden layers, and four neurons

Decision Tree (DTs)

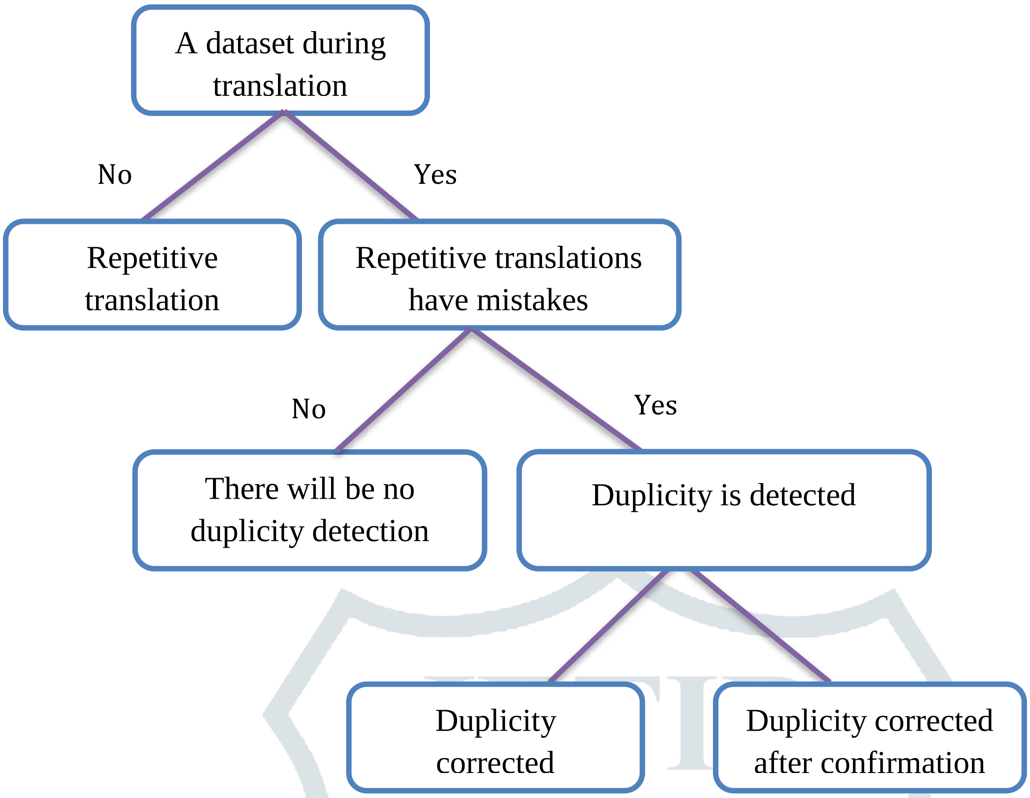

Figure 5 A set of fundamental decision trees for the management and data [32]

Decision trees (DTs) (CNNs) are significantly less complex in logic than convolutional neural networks (CNNs). A goal-setting approach is built using a decision tree topology, which is a unique flowchart or network of options (nodes) and likely effects (divisions or leaves) [25, 37]. The process is repeated recursively on each dependent subdivision set up until there is no longer any room for division or a distinct categorization can be determined. Decision trees have many benefits, including being simple to comprehend and examine, requiring little data scheduling, handling a wide range of data types, including present value (named), computation, and categorical variables, focusing on producing rigorous classifiers, those that are quick to "peruse," and being verifiable using the normality test. In contrast, DTs frequently perform worse than CNNs in increasingly challenging classification tasks [38].

Support Vector Machines (SVM)

Support Vector Machines employ supervised learning techniques; consequently they require labeled, previously studied data to categorize newly discovered data. The fundamental strategy for categorizing the data begins with an effort to develop a mechanism that divides the information points into the right labels with (a) the lowest number of mistakes or (b) the widest gap. Due to the labels' improved ability to identify each other, greater vacant spaces near to the division function produce less mistakes.

Figure 7 shows that a collection of data may very easily be separated by numerous variables without making any mistakes. As a result, an extra variable is utilized to assess the separation's excellence: the space

surrounding the separation functional. Because it more clearly separates the two lessons, separation A is preferable in this situation.

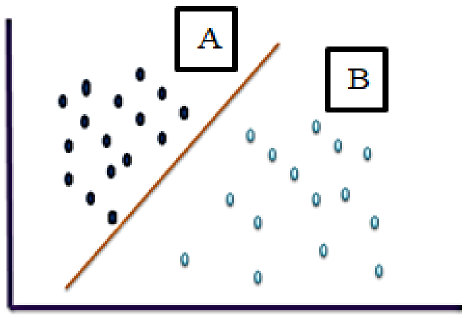

Figure 6 Visualization of a Support Vector Machine splitting a data set into two classes

One or more hyperplanes may be created by Support Vector Machines in a space with n dimensions. The method of dividing the data always starts with a first effort to try to linear divide the information into the matching categories. A data collection of n data points is used in the illustration's job of estimating the likelihood that a client would make an investment in an online store. Each information point is composed of a label (y2f purchase; no purchase g) and an attribute vector (x) holding the data values for those particular sessions. The Support Vector Machine is currently attempting to identify an equation that distinguishes between every point of data of the type (x; y) with y = yes and all data values of the form (x; y) with y = no.

Results and discussion

The following section gives the numbers to the approach described in previous chapter. Table 1 gives an overview over the amount of sentences, attributes and documents used for the given tasks.

Table 1 Stats of the datasets utilized

Count of machine translation systems	5
The quantity of scientific papers	15
Sentence counts in all original documents	40, 000
Phrase counts in the data sets, each translation method	25, 000
About how many sentences are in the sets of data	45, 500
Amount of data sets overall	10
The quantity of characteristics that require a reference translation	14
Quantity of qualities without reference translation	20
Quantity of every attribute	35
Maximum number of variables for each characteristic	1, 531, 000
Number of fabricated documents produced	20, 280

Table 2 lists the nine data sets that were utilized, with back and forth translates for six out of the nine data sets generated with the no-cost translation tool and the five automated translation systems as contenders and sources, accordingly.

Table 2 Utilized a mixture of recommendations and applicants

Candidate	Source1	Source2	Back and forth Reference
QuillBoat	Bing	—	QillBoat RTT via Freetranslation
Google Translate	Freetranslation	—	Google RTT via Freetranslation
Editab	Bing	Google	Google RTT via Freetranslation

Grammarly	Google	—	Google RTT via Freetranslation
Bing Translator	Freetranslation	—	Editab RTT via Freetranslation
QuillBoat	Google	Editab	Bing RTT via Freetranslation
Freetranslation	Editab	—	—
Editab	Bing	—	—
Grammarly	Google	Bing	—

By performing the procedure of optimization for various settings while turning on and off certain processes, like in the instance provided, it was possible to assess the impact of the procedures listed previously. Duplicates = true, Normalize Data = false, Detect Outlier = true, Remove Related Attributes = false.

Table 3 A summary of the designs and optimizations that were produced

Method	**Research Question 1**
Decision Trees Method	490, 000
Neural Networks Method	6, 500
k-Nearest Neighbors Method	3, 720
Support Vector Machines Method	2, 750
Total Number of improvement	502, 970

Table 4 aggregates the data that will be provided in more depth later on and indicates the overall effectiveness of the employed algorithms so as to give a concise summary.

Table 4 Calculated algorithm means and standard deviations

Method	**Precision**	**Standard Deviation**
Decision Tree	69.72%	0.015
Artificial Neural Network	**72.67%**	0.015
k-Nearest Neighbor	70.46%	**0.010**
Naive Bayes	66.98%	0.020
Support Vector Machines	62.54%	0.018

Table 5 A summaries of the most effective Decision Tree outcomes for the corresponding candidate-reference pairings

Decision Tree					
Candidate	Source1	Source2	Accuracy	F1-Automated	F1-Professional
QuillBoat	Bing	—	70.01%	0.735	0.650
Google Translate	Freetranslation	—	68.11%	0.716	0.647
Editab	Bing	Google	71.76%	0.755	0.652
Grammarly	Google	—	71.32%	0.716	0.699
Bing Translator	Freetranslation	—	69.45%	0.698	0.685
QuillBoat	Google	Editab	71.33%	0.725	0.691
Freetranslation	Editab	—	67.45%	0.699	0.631
Editab	Bing	—	68.96%	0.697	0.670

Grammarly	Google	Bing	68.63%	0.696	0.676

The findings for phrase forecasts utilizing Decision Trees, together with the corresponding F1-scores, are shown in Table 5. A total of 50,000 Decision Trees were created and assessed for each candidate-reference pair given in the table. The utilized data comprised all accessible characteristics adjusted, and its size was further decreased by up to 5% by a single outlier identification and the elimination of duplication.

Table 6 Outline of the most effective candidate-reference pairings for artificial neural networks

Artificial Neural Network					
Candidate	Source1	Source2	Precision	F1-Automated	F1-Professional
QuillBoat	Bing	—	70.16%	0.751	0.667
Google Translate	Freetranslation	—	69.28%	0.718	0.658
Editab	Bing	Google	72.99%	0.757	0.674
Grammarly	Google	—	**73.87%**	0.738	0.726
Bing Translator	Freetranslation	—	70.76%	0.701	0.698
QuillBoat	Google	Editab	72.33%	0.735	0.721

Table 7 displays the outcomes for the employed k-Nearest Neighbor method. 50 simulations were created and assessed for every candidate-reference pairing. The outcomes show which models had the greatest obtained accuracy. Similarly to the previously discussed configurations, the data is prepared, and the set of data decreases by up to 5% outliers and owing to the elimination of copies, while the characteristics are normalized in order to execute the outlier identification.

Table 7 presents an overview of the top k-Nearest Neighbor outcomes for the various candidate-reference pairings

k-Nearest Neighbor					
Candidate	Source1	Source2	Precision	F1-Automated	F1-Professional
QuillBoat	Bing	—	71.14%	0.691	0.684
Google Translate	Freetranslation	—	68.27%	0.707	0.697
Editab	Bing	Google	71.92%	0.718	0.688
Grammarly	Google	—	**73.89%**	0.714	0.691
Bing Translator	Freetranslation	—	70.75%	0.704	0.687
QuillBoat	Google	Editab	72.37%	0.719	0.692

The outcomes of assessments conducted with the method known as Naive Bayes are shown in the Table 8. The findings are definitive for the information that was utilized set and cannot be improved optimized due to the limitations of the technique. The data set is comparable to the preceding findings: All readily available characteristics are used and standardized. The data set is duplicate-free and has outlier reduction of up to 5%.

Table 8 A summary of the findings of the Naive Bayes models for the corresponding candidate-reference pairings

Naive Bayes					
Candidate	Source1	Source2	Accuracy	F1-Automated	F1-Professional
QuillBoat	Bing	—	68.56%	0.700	0.653
Google Translate	Freetranslation	—	69.10%	0.703	0.641
Editab	Bing	Google	**70.34%**	0.730	0.666
Grammarly	Google	—	64.33%	0.682	0.593
Bing Translator	Freetranslation	—	66.56%	0.691	0.610
QuillBoat	Google	Editab	65.32%	0.690	0.580

Table 9 displays the Support Vector Machines' greatest accuracy results. The data set and properties employed are comparable to the previously demonstrated sentence-based methods. For 50 iterations, SVM have been improved.

Table 9 A summaries of the most effective SVM findings for the corresponding candidate-reference pairings

SVM					
Candidate	Source1	Source2	Precision	F1-Automated	F1-Professional
QuillBoat	Bing	—	**64.66%**	0.615	0.666
Google Translate	Freetranslation	—	61.12%	0.581	0.628
Editab	Bing	Google	63.78%	0.638	0.637
Grammarly	Google	—	60.42%	0.650	0.547
Bing Translator	Freetranslation	—	60.50%	0.578	0.628
QuillBoat	Google	Editab	64.34%	0.669	0.560

Conclusions

Given that the source text is available, why can the translated level of technical publications are assessed? This was accomplished by employing an established finding method that involves the steps of data collection, preliminary processing, selection of a suitable data mining technique, and interpretation of the trends discovered. The findings were then applied to more study. The sentence-level breakdown of the written material database resulted in nine sets of data, each of which has 22,327 data items for the two different interpretation types (automatic interpretation and expert translation). It was decided to use 32 measurements and characteristics, of which 18 required an initial language for the computation procedure and 14 required not. One or two automated translation methods were utilized for translating the source text and provide references for the provided candidate texts in order to construct a standard translation. The presented data set underwent preprocessing, including normalization of the data to produce similar values for attributes, removal of 5% anomalies, and feature correlations of higher than 90%. Five artificial intelligence algorithms-Decision Trees, k-Nearest Neighbor, Artificial Neural Networks, Naive Bayes, and SVM were applied in several iterations with the preprocessed dataset. After learning the algorithms, the methods' variables were tuned and they were evaluated on an exception set that was extracted from the database. The k-Nearest Neighbor classification produced the best outcomes, scoring 73 when given knowledge to the original file and 63% when not. The optimised methods were applied to the words of each initial record to make a claim about paragraph level categories, and the outcomes were then re-combined to categorize the relevant documents.

In addition, a structure for rating phrases and papers according to their excellence, irrespective of the kind of translation, was developed. The proposed approach divides phrases into four groups using two improved artificial intelligence designs, and it also includes a reference-independent spelling and grammatical checker that creates an average error score for each phrase. The sentence quality categories of the relevant phrases are summed up, with greater weight given to sentences with a higher error rate.

References

1. H. Somers, "Round-trip translation: what is it good for," in In proceedings of the Australasian Language Technology Workshop, pp. 127–133, 2005.

2. K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "Bleu: a method for automatic evaluation of machine translation," in Computational Linguistics (ACL), Proceed- ings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, pp. 311–318, 2002.

3. G. Doddington, "Automatic evaluation of machine translation quality using n-gram co-occurrence statistics," in Proceedings of the Second International Conference on Human Language Technology Research, HLT '02, (San Francisco, CA, USA), pp. 138–145, Morgan Kaufmann Publishers Inc., 2002.

4. A. Lavie and A. Agarwal, "Meteor: An automatic metric for mt evaluation with improved correlation with human judgments," pp. 65–72, 2005.

5. A. Kulesza and S. M. Shieber, "A learning approach to improving sentence-level mt evaluation," in In Proceeding of the 10th International Conference on Theoretical and Methodological Issues in Machine Translation, 2004.

6. M. Popovic´, D. Vilar, E. Avramidis, and A. Burchardt, "Evaluation without refer- ences: Ibm1 scores as evaluation metrics," 2011.

7. M. Gamon, A. Aue, and M. Smets, "Sentence-level mt evaluation without reference translations: Beyond language modeling," in In European Association for Machine Translation (EAMT), 2005.

8. J. S. Albrecht and R. Hwa, "2007b. regression for sentence-level mt evaluation with pseudo references," in In Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics (ACL), pp. 296–303, 2007.

9. J. S. Albrecht and R. Hwa, "The role of pseudo references in mt evaluation," in In Proceedings of ACL, 2008.

10. U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From data mining to knowledge discovery in databases," Ai Magazine, vol. 17, pp. 37–54, 1996. 11. J. H. Friedman, "Data mining and statistics: What's the connection," in Proceedings of the 29th Symposium on the Interface Between Computer Science and Statistics, 1997.

11. O. Maimon and L. Rokach, eds., Data Mining and Knowledge Discovery Handbook, 2nd ed. Springer, 2010.

12. L. Rokach and O. Z. Maimon, Data mining with decision trees: theory and appli- cations, vol. 69 of Series in machine perception and artificial intelligence. World Scientific Publishing Co., 2008.

13. B. P. Battula and R. S. Prasad, “An Overview of Recent Machine Learning Strate- gies in Data Mining,” International Journal of Advanced Computer Science and Applications, vol. 4, 2013.

14. 8. Simes RJ (1985) Treatment selection for cancer patients: application of statistical decision theory to the treatment of advanced ovarian cancer. J Chronic Dis 38:171-186.

15. 9. Maclin PS, Dempsey J, Brooks J, Rand J 1991 Using neural networks to diagnose cancer. Journal of medical systems 15(1):11-19.

16. 10. Cicchetti DV (1992) Neural networks and diagnosis in the clinical laboratory: state of the art. Clin Chem 38:9-10.

17. 11. Petricoin EF, Liotta LA (2004) SELDI-TOF-based serum proteomic pattern diagnostics for early detection of cancer. Curr Opin Biotechnol 15:24-30.

18. Bocchi L, Coppini G, Nori J, Valli G (2004) Detection of single and clustered microcalcifications in mammograms using fractals models and neural networks. Med Eng Phys 26:303-312.

19. Zhou X, Liu KY, Wong ST (2004) Cancer classification and prediction using logistic regression with Bayesian gene selection. J Biomed Inform 37:249-259.

20. Dettling M (2004) BagBoosting for tumor classification with gene expression data. Bioinformatics 20:3583-593.

21. Wang JX, Zhang B, Yu JK, Liu J, Yang MQ, Zheng S (2005) Application of serum protein fingerprinting coupled with artificial neural network model in diagnosis of hepatocellular carcinoma. Chinese medical journal 118(15):1278-1284.

22. Mccarthy JF, Marx KA, Hoffman PE, Gee AG, O'neil P, Ujwal ML, Hotchkiss J (2004) Applications of machine learning and high-dimensional visualization in cancer detection, diagnosis, and management. Ann N Y Acad Sci 1020:239-262.

23. Hagerty RG, Butow PN, Ellis PM, Dimitry S, Tattersall MH (2005) Communicating prognosis in cancer care: a systematic review of the literature. Ann Oncol 16:1005-1053.

24. Mitchell T (1997) Machine Learning. New York: McGraw Hill.

25. Buch VH, Ahmed I, Maruthappu M (2018) Artificial intelligence in medicine: Current trends and future possibilities. Br J Gen Pract 68:143-144.

26. Liu Y, Chen PH, Krause J, Peng L (2019) How to read articles that use machine learning: Users’ guides to the medical literature. JAMA 322:1806-1816.

27. Csaba Szepesvari (2010) Algorithms for Reinforcement Learning: Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool publishers, Williston. 4(1):1-103. (https://doi.org/10.2200/S00268ED1V01Y201005AIM009)

28. Adel Aloraini (2012) Different machine learning algorithms for breast cancer diagnosis. (IJAIA) 3(6):21.

29. Sivakami K, Saraswathi N (2015) Mining big data: breast cancer prediction using DT-SVM hybrid model. IJSEAS 5:418-429.

30. Megha Rathi , Arun Kumar Singh (2012) Breast Cancer Prediction using Naïve Bayes Classifie, Int J Inf Technol 1(2):77-80.

31. Cuong Nguyen, Yong Wang, Ha Nam Nguyen (2013) Random forest classifier combined with feature selection for breast cancer diagnosis and prognostic. J Biomed Sci and Eng 6:551- 560.

32. Hinton GE (1992) How neural networks learn from experience. Sci Am 267:144-151.

33. Furundzic D, Djordjevic M, Bekic AJ (1998) Neural networks approach to early breast cancer detection. J Syst Archit 44:617–633.

34. Minsky M, Papert S (1969) Perceptrons; MIT Press: Cambridge, MA, USA.

35. Rumelhart DE, Mcclellend JL (1986) Parallel distributed processing: explorations in the microstructure of cognition; MIT Press: Cambridge, MA, USA.

36. Quinlan JR (1986) Induction of decision trees. Machine Learning 1:81-106.

37. Atlas LE, Cole RA, Connor JT, El-Sharkawi MA, Marks II RJ, Muthusamy YK, Barnard E (1990) Performance comparisons between backpropagation networks and classification trees on three real-world applications. NeurIPS 2:622-629.

38. I. Steinwart and A. Christmann, Support Vector Machines. Springer Publishing Company, Incorporated, 1st ed., 2008.

39. D. M. W. Powers, "Evaluation: From Precision, Recall and F-Factor to ROC, In- formedness, Markedness & Correlation," Tech. Rep. SIE-07-001, School of Infor- matics and Engineering, Flinders University, Adelaide, Australia, 2007.