Iyer, S. et al. (Eds.) (2022). Proceedings of the 30th International Conference on Computers in Education.
Asia-Pacific Society for Computers in Education

Innovative and Effective Spreadsheet Tool for Learning Sentiment Analysis and Prediction

Michelle CHEONG[a*] & Jean CHEN[a]
[a]*School of Computing & Information Systems, Singapore Management University, Singapore*
*michcheong@smu.edu.sg

Abstract: In this paper, we describe a spreadsheet tool which was developed and implemented to teach sentiment analysis and prediction to university students who are less technically inclined. We adopted the "Use-Modify-Create" cycle to ensure that the students learn the computing concepts and practices in a progressive and active learning manner, and we evaluated the effectiveness of the tool using the Learning Object Evaluation Scale for Students (LOES-S). The evaluation shows that our tool performed well in terms of quality, learn and engagement constructs, and our paired T-test shows strong evidence that, on average, the course module with the tool does lead to improvement in the students' knowledge in sentiment analysis and prediction.

Keywords: Sentiment analysis and prediction, spreadsheet tool, less technically inclined students

1. Introduction

Teaching and learning technical topics such as sentiment analysis and prediction can be daunting for the instructors and students, especially for students who are less technically inclined. To achieve effective teaching and learning, digital education system needs to incorporate technologies and tools to assist the teaching and learning process (Williamson, 2016).

One such tool is the Text Sentiment Analysis Excel Add-in from the Azure Machine Learning Web Services available in the MS Office Store. However, such an Add-in will only function as a black box to the users and using it to teach students will not be able to achieve the intended learning outcomes such as, interpreting the theory and concepts of sentiment analysis and prediction, changing the parameters to improve the prediction results, analyzing the prediction results obtained, and deploying in new implementations. As such, the "Use-Modify-Create" cycle (Lytle, 2019) which is commonly used to learn computing concepts and practices in a progressive and active learning manner, cannot be implemented. Expecting the less technically inclined students to develop their own sentiment analysis and prediction tool using R or Python programming would be challenging, and the intended learning outcomes may be overshadowed by the technical difficulties met during coding, leading to students 'missing the forest for the trees'.

In this study, we explain how we used an innovative spreadsheet tool in a classroom for the learning of sentiment analysis and prediction, following the "Use-Modify-Create" cycle. In addition, we will provide a statistical analysis of the students' knowledge gain based on pre-test and post-test results and share the students' quantitative and qualitative feedback on using the tool. The results of this study can inspire educators to consider using easy-to-use and effective tool to teach technical topics to students who are less technically inclined, to achieve the intended learning outcomes.

2. Literature Review

To prepare the next generation of learners to have a basic understanding of machine learning, many educators have designed tools to achieve this objective. von Wangenhelm et al. (2021) provided a ten-year systematic mapping of visual tools for teaching machine learning (ML) in K-12. 16 tools were identified, and most of them support both the development of ML models and the deployment as software artifacts. The deployment platforms are mainly block-based programming environments including Scratch (Agassi et al., 2019; Alturayeif et al., 2020; Druga et al,, 2019; Zimmermann-Niefield et al., 2020), App Inventor (Tang et al., 2019; Zhu, 2019), and Snap! (Jatzlau et al., 2019; Kahn et al., 2018). The application areas are mainly in image recognition, speech recognition and synthesis, and object detection.

Batrinca & Treleaven (2015) surveyed techniques, tools and platforms for social media analytics. In the paper, they reviewed software tools and explained how to use them to scrape, cleanse and analyze social media data, mainly to benefit researchers to carry out their research works, rather than to use them for teaching and learning. Specifically for text analysis tools, they listed Python NLTK – Natural Language Toolkit (www.nltk.org) and GATE (http://gate.ac.uk/sentiment). While both NLTK and GATE have been used for teaching in a few institutions, they will require the students to have software engineering knowledge and programming skills in Python or Java.

R is another programming language that is often used to perform word cloud and sentiment analysis. Kabir, Ahmed & Karim (2020) described how they used R programming language and R Studio IDE to analyze Amazon earphones online reviews to know people's attitude, opinions, reviews and sentiments. They discussed their research steps in terms of how they read in the data set to build the corpus, perform data cleaning, create the term document matrix, create the bar plot and word clouds, obtain the sentiment score, and finally draw conclusions on the opinions and sentiments for the earphones. This work is a suitable reference for anyone to replicate what they have done but was not designed as a teaching material to be delivered in a classroom.

The work by Subramanian & Cote (2018) is the closest match to our study. In their paper, they described how a fully online graduate course in text mining using SAS Enterprise Miner and Twitter data were taught at the Quinnipiac University. They listed the topics taught in the course and shared the details of the student's project in terms of the data, deliverables expected, and grading rubrics. One selected student's report was presented as appendix. While they concluded that the course was a success, there was no data collected and analysed to evaluate how effective the course was, and no evaluation metric was used to assess whether SAS Enterprise Miner was indeed an effective tool to use in the teaching and learning.

To assess the effectiveness the use of software tools for learning, evaluation metrics for learning objects can be applied. IEEE Learning Technology Standards Committee (LTSC) defined learning object as "any entity, digital or non-digital, which can be used, re-used or referenced during technology-supported learning." (LTSC, 2007). Kay & Knaack (2008) proposed a multi-component model, known as the Learning Object Evaluation Metric (LOEM) with four constructs: interactivity, design, engagement, and usability, based on comprehensive review of literature. They tested the model with 1113 students, 33 teachers and 44 learning objects, and showed good internal and inter-rater reliability, and all four constructs correlated significantly with student learning performance. The same study also developed a reliable student-based evaluation tool for assessing learning objects, known as the Learning Object Evaluation Scale for Students (LOES-S) using three constructs: learning, quality, and engagement (Kay & Knaack, 2009).

This study aims to explain how the tool was implemented in a classroom for learning sentiment analysis and prediction following the "Use-Modify-Create" cycle, for less technically inclined university students. We provide a statistical analysis of the students' knowledge gain based on pre-test and post-test results. In addition, we adopted the LOES-S proposed by Kay & Knaack (2009) to perform the evaluation of tool effectiveness from the students' perspective.

3. Sentiment Analysis and Prediction Tool

3.1 Text Pre-Processing and Sentiment Prediction

To perform sentiment analysis and prediction, the first step is text pre-processing which includes tokenization and normalization, where normalization will involve stemming and lemmatization. In stemming, part of the word will be cut off, either at the beginning or the end of the word, taking into account a list of common prefixes and suffixes that can be found in an inflected word. For example, "studies" will be stemmed into "studi", while "studying" will be stemmed into "study". Such indiscriminate and crude way of cutting will not always result in words which are easily recognizable. Lemmatization attempts to overcome the shortcomings of stemming, by considering the full morphological analysis to identify the lemma for each word. For example, “studies” and “studying” will both be linked to the lemma “study”. To do so, it is necessary to have detailed dictionaries which the algorithm can look through to link the inflected word back to its lemma.

Due to the shortcomings in stemming and the complexity in lemmatization, the tool that was developed does not include both features. Instead, it will include stemming to remove unnecessary characters that will obscure the word. For example, “…apple” will be front stemmed into “apple” with “…” cut away, while “orange…” will be back stemmed into “orange” with “…” cut away. The objective is to ensure that complete words are kept in the dictionary for easy understanding and analysis.

The Naïve Bayes’ Classifier is implemented in the tool to classify a new text into positive or negative sentiment. Neutral sentiment was not included in the prediction, and other classifiers can also be implemented easily in future versions of the tool.

3.2 The Tool

The tool was developed using Excel and VBA programming and the workbook consists of nine sheets. Excel was used as it is a tool most people are comfortable with, and the buttons with VBA programs assigned will improve usability tremendously. The first sheet **PasteHere** is the main user interface where the user will paste the labeled texts to train the machine learning algorithm, as well as to paste new texts for sentiment prediction. At the top of the **PasteHere** sheet are eight buttons, and each button has a VBA macro program assigned, for the user to perform automated tasks and computations. Figure 1 shows the user interface on **PasteHere** sheet with the eight buttons, and the arrows added between the buttons will guide the user how to use the tool in a sequential manner.

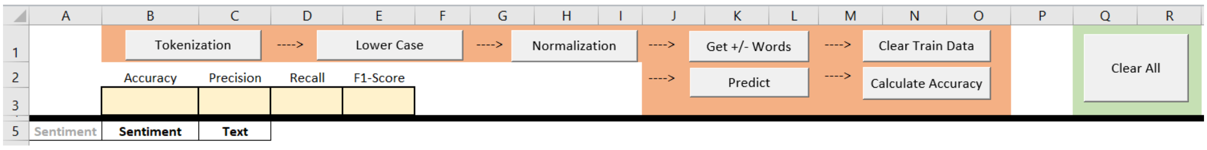

Figure 1. User Interface of PasteHere sheet.

The eight other sheets as described below:

- **Work** sheet – This is an empty sheet initially. It will be filled with useful words which are left over after the VBA program performs word removal and word stemming.
- **Positive** sheet – This is an empty sheet initially. It will be filled with words which are extracted from positive texts in the training data set. Each word will have its associated frequency count, log probability and conditional probability.
- **Negative** sheet – This is an empty sheet initially. It will be filled with words which are extracted from negative texts in the training data set. Each word will have its associated frequency count, log probability and conditional probability.
- **Stopwords** sheet – This sheet contains a list of 174 stop words (e.g. a, is, of, the) retrieved from http://www.ranks.nl/stopwords.
- **FrontRemove** sheet – This sheet contains a list of prefixes (e.g. #, @, http, RT) which the VBA program will remove words with these prefixes.
- **Otherwords** sheet – This sheet contains a list of words which are neither Stopwords nor FrontRemove words. It is useful for words which are specific to the data set and the user would like to have such words removed.
- **FrontStems** sheet – This sheet contains a list of special characters which the user would like to have them stemmed from the front of the words.
- **BackStems** sheet – This sheet contains a list of special characters which the user would like to have them stemmed from the back of the words.

The tool is flexible to accommodate any text for text pre-processing and sentiment prediction, as the user can edit the work lists in the **FrontRemove**, **Otherwords**, **FrontStems** and **BackStems** sheets.

4. Research Methodology and Implementation

4.1 Research Methodology

The tool was used in a course module "COMM629 Analytics" with 38 students pursuing the Master of Communications Management programme offered by the Lee Kong Chian School of Business at the Singapore Management University. The module was delivered over a 5-week period with a 3-hour class session each week. The topic on sentiment analytics was delivered in the third session. During this session, theory and concepts related to sentiment analytics were taught, before a hands-on laboratory session was conducted to guide the students to "Use" the tool for sentiment analysis and prediction for a set of Twitter data. In the "Modify" step, the students will edit the list of words in the **FrontRemove**, **Otherwords**, **FrontStems** and **BacksStems** sheets to learn how they can affect the creation of positive and negative word features, which will in turn affect the prediction accuracy. In the "Create" step, students are tasked to perform a new experiment using self-sourced data set, and apply what they have learned in the "Modify" step to obtain as high a prediction accuracy as possible for the new data set.

4.2 Classroom Exercise

The class exercise consists of three main parts and students only need to click the buttons in sequence as explained in section 3.3:

- Part 1 is to use the training data with labeled tweets to obtain the positive and negative words with their respective frequency count, log probability and conditional probability computed.
- Part 2 is to use the test data set as new data for the machine to predict the sentiment for each new tweet.
- Part 3 is to compute the accuracy, precision, recall and F1-score by comparing the machine predicted sentiment with the user labeled sentiment for the test data set.

4.3 Pre-Test and Post-Test

The students took the pre-test at the start of session 1, after they have given consent to take part in the study, and they took the post-test at the end of session 5. There were ten Multiple-Choice Questions and for each question, four answer options were provided where the last answer option is always "I do not know the answer". The students were advised to choose the last answer option "I do not know the answer" if they are not sure of the answer instead of making a wild guess. This is to prevent the situation where students managed to obtain the correct answer due to random selection, which will affect the computation of actual knowledge gain.

4.4 Feedback Survey and Evaluation Metric

As there is no control group (where students will not get to use the tool), a survey was conducted at the end of Session 5 with the students to collect quantitative and qualitative feedback on the effectiveness of the tool. The four survey questions match the three constructs: learn, quality, and engagement in the Learning Object Evaluation Scale for Students (LOES-S), proposed by Kay & Knaack (2009):

- Q1: The tool is easy to understand and use. (Quality)
- Q2: I learn how to use the tool to pre-process text to create word features for positive and negative sentiment prediction. (Learn)
- Q3: By using the tool, I can understand and apply the concepts and principles behind sentiment analytics better. (Learn and Engagement)
- Q4: By using the tool, I can easily conduct experiments on sentiment analytics. (Engagement)

5. Effectiveness of the Tool for Learning Sentiment Analysis and Prediction

5.1 Pre-Test and Post-Test Analysis

With the pre-test and post-test average scores, we conducted a paired T-test to test the null hypothesis (H0) that the true mean difference between the scores is zero, and obtained a p-value of <0.0001 (<

0.05, reject H0) and t-statistic of 13.092. Thus, there is strong evidence that, on average, the course module with the tool does lead to improvement in the students' knowledge in sentiment analysis and prediction. It is acknowledged that there is no control group to make the comparison in knowledge gain, as there is no alternative course that teaches the same topic without the use of the tool.

5.2 Quantitative Feedback Analysis

To gain more insights into the effectiveness of the tool in learning sentiment analysis and prediction, a survey with four questions, each with five options ("Strongly Disagree", "Disagree", "Neutral", "Agree", and "Strongly Agree") was conducted. Based on the percentage of students who rated "Agree" and "Strongly Agree", the feedback results in Table 1 show that the tool performed very well. The best construct was Learn (78.4%), followed by Learn and Engagement (75.0%), then Quality (67.6%). Engagement in terms of using the tool to conduct experiments scored the lowest (58.3%). A plausible reason could be that the students needed more practice before they could conduct new experiments independently. Recall that the topic on sentiment analytics was delivered only in one 3-hour session, which could be insufficient to cover the theory and concepts, and to conduct the hands-on laboratory on using the tool. Providing more time and more practice exercises will improve this construct in future.

Table 1. *Quantitative Feedback Results*

	% Rated Agree and Strongly Agree
Q1: The tool is easy to understand and use. (Quality)	67.6%
Q2: I learn how to use the tool to pre-process text to create word features for positive and negative sentiment prediction. (Learn)	78.4%
Q3: By using the tool, I can understand and apply the concepts and principles behind sentiment analytics better. (Learn and Engagement)	75.0%
Q4: By using the tool, I can easily conduct experiments on sentiment analytics. (Engagement)	58.3%

5.3 Qualitative Feedback Analysis

Out of 38 students, 15 students did not provide qualitative feedback. Of the 23 feedback comments, almost all of them are positive which include the following:

- "The tool enables us to approach data and its analysis in methodological way, without involving complex equations."
- "Through the simple-to-understand steps, the learning and pickup of skills are no-issue for any layman."
- "The tool strengths my understanding of the data analysis, the background and the meaning."
- "Easy to use and helps in understanding of certain fundamental steps necessary to be conducted before sentiment analysis can be done."

Other feedback comments include "Need more practice to understand the usefulness" and "Can it be enhanced to include neutral sentiment", which are the potential improvement areas for future implementations.

6. Conclusions

In the age of digital education, learning technologies will enhance students' learning immensely. The challenging task in teaching technical topics to the less technically inclined students will require tools which are easy-to-use and yet effective. Our spreadsheet tool for the learning of sentiment analysis and prediction was shown to perform very well in terms of quality, learn and engagement constructs, in a course module delivered to 38 business school university students. The statistical analysis of the pre-test and post-test scores shows strong evidence that, on average, the course module with the tool does lead to improvement in the students' knowledge in sentiment analysis and prediction. Using LOES-S to perform the evaluation of tool effectiveness from the students' perspective, the tool performed very well in Learn and Engagement constructs. Future improvements include adding in neutral sentiment prediction, prediction using a few other classifiers for comparison, and allowing students more practice opportunities to enhance learning.

Acknowledgements

Due to the need to collect data on students who are participants in the study, IRB application was submitted and approved as "Category 1:Exempt from Further IRB Review" with IRB Exemption Number IEB-22-0430E023(422).

References

Agassi, A., Erel, H., Wald, I. Y., & Zuckerman, O. (2019). Scratch nodes ML: A playful system for children to create gesture recognition classifiers. *Proceedings of the Conference on Human Factors in Computing Systems*, ACM, 1–6. https://doi.org/10.1145/3290607.3312894

Alturayeif, N., Alturaief, N., & Alhathloul, Z. (2020). DeepScratch: Scratch programming language extension for deep learning education. *International Journal of Advanced Computer Science and Applications*, 11(7), 642–650. https://doi.org/10.14569/IJACSA.2020.0110777

Batrinca, B., & Treleaven, P.C. (2015). Social media analytics: a survey of techniques, tools and platforms. *AI & Soc*, 30, 89–116. DOI 10.1007/s00146-014-0549-4

Druga, S. Vu, S. T., Likhith, E., & Qiu T. (2019). Inclusive AI literacy for kids around the world. *Proceedings of Fab Learning*, ACM 104–111. https://doi.org/10.1145/3311890.3311904

Jatzlau, S. Michaeli, T., Seegerer, S., & Romeike, R (2019). It's not Magic After All – Machine Learning in Snap! using Reinforcement Learning. *Proceedings of IEEE Blocks and Beyond Workshop*, Memphis, TN, USA, 37–41. https://doi.org/10.1109/BB48857.2019.8941208

Kabir, A.I., Ahmed, K., & Karim, R. (2020). Word Cloud and Sentiment Analysis of Amazon Earphones Reviews with R Programming Language. *Informatica Econo*mica, 20(4), 55–71

Kay, R. H., & Knaack, L. (2008). A multi-component model for assessing learning objects: The learning object evaluation metric (LOEM). *Australasian Journal of Educational Technology*, 24(5), 574–591

Kay, R. H., & Knaack, L. (2009). Assessing learning, quality and engagement in learning objects: the Learning Object Evaluation Scale for Students (LOES-S). *Education Technology Research and Development*, 57, 147–168. DOI 10.1007/s11423-008-9094-5

Kahn, K. M., Megasari, R., Piantari, E., & Junaeti, E. (2018). AI Programming by Children Using Snap! Block Programming in a Developing Country. In *Proceedings of the 13th European Conference on Technology Enhanced Learning*, Leeds, UK.

Learning Technology Standards Committee (LTSC). (2007). The Learning Object Metadata standard. IEEE.

Lytle, N., et al. (2019). Use, Modify, Create: Comparing Computational Thinking Lesson Progressions for STEM Classes. *Proceedings of the Conference on Innovation and Technology in Computer Science Education*, ACM, 395–401. https://doi.org/10.1145/3304221.3319786

Subramanian, R., & Cote, D. (2018). Using SAS™ Software to Enhance Pedagogy for Text Mining and Sentiment Analysis Using Social Media (Twitter™) Data. *Journal of International Technology and Information Management*, 27(2), 73–98.

Tang, D., Utsumi, Y., & Lao, N. (2019). PIC: A Personal Image Classification Webtool for High School Students. In *Proceedings of the IJCAI EduAI Workshop*, Macao, China.

von Wangenheim, C.G., Hauck, J.C.R., Pacheco, F.S., & Bueno, M.F.B. (2021). Visual tools for teaching machine learning in K-12:A ten-year systematic mapping. *Education and Information Technologies*, 26, 5733–5778. https://doi.org/10.1007/s10639-021-10570-8

Williamson, B. (2016). Digital education governance: data visualization, predictive analytics, and 'real-time' policy instruments. *Journal of Education Policy*, 31(2), 123–141.

Zhu, K. (2019). An Educational Approach to Machine Learning with Mobile Applications. M.Eng thesis, MIT, Cambridge, MA, USA.

Zimmermann-Niefield, A., Polson, S., Moreno, C., & Shapiro, R. B. (2020). Youth making machine learning models for gesture-controlled interactive media. In *Proceedings of the Interaction Design and Children Conference*. ACM, 63–74. https://doi.org/10.1145/3392063.3394438