Towards Ensemble Learning for Tracking Food Insecurity From News Articles

Andrew Lukyamuzi, Mbarara University of Science and Technology, Uganda

John Ngubiri, Makerere University, Uganda

Washington Okori, Uganda Technology and Management University, Uganda

ABSTRACT

The study integrates ensemble learning into a task of classifying if a news article is on food insecurity or not. Similarity algorithms were exploited to imitate human cognition, an innovation to enhance performance. Four out of six classifiers generated performance improvement with the innovation. Articles on food insecurity identified with best classifier were generated into trends which were comparable with official trends. This paper provides information useful to stake holders in taking appropriate action depending on prevailing conditions of food insecurity. Two suggestions are put forth to promote performance: (1) using articles aggregated from several news media and (2) blending more classifiers in an ensemble.

KEYWORDS

Bag-of-Words, Deep Learning, Human Cognition, Text Classification, Text Similarity Measure, Traditional Machine Learning

INTRODUCTION

Ensemble Learning, the process of blending two or more algorithms to promote performance is now at the frontier in field of Machine Learning. The Netflix winner of 1 million dollar prize used an ensemble learning of over 100 models (Andreas, Jahrer, Bell, & Park, 2009). The winner had been required to improve accuracy by 10% of Netflix's system for user movie recommendations (Hallinan & Striphas, 2016). The challenge attracted over 50,000 contestants from 186 countries (NETFLIX, 2009). The second best contestant also employed Ensemble Learning (Feuerverger, He, & Khatri, 2012). In the Kaggle competition of 2015, the winner for forecasting six weeks of store sales used ensemble learning. Using an ensemble of residual nets, He, Zhang, Ren, & Sun (2015) won 1st place on the ILSVRC 2015 classification task. To emphasize further the power of ensemble learning, Wind & Winther (2014) assert that Kaggle competitions are frequently won by competitors who integrate some aspect of ensemble learning. The power of ensemble learning/blended approach has been exhibited in other human endeavors (for example see Abdel Aziz et al., 2016; Bouzaida & Sakly, 2018; Hussein et al., 2017; Majhi, 2018). From this, it can be inferred that high-end performance is highly probable with ensemble learning.

DOI: 10.4018/IJSDA.2020100107

Deep learning has also come on the stage in comparison to traditional Machine Learning. A distinguishing feature of Deep Learning is the ability to automatically extract features as opposed to manual feature engineering as it is in traditional Machine Learning. Deep Learning uses multiple layers of non-linear operations to offer state-of-the-art systems such as in vision and language tasks (Ren, Zhang, & Suganthan 2016). Because of its tendencies to generate better performance than traditional Machine Learning, deep learning is increasingly becoming preferable (W. Li, Liu, Zhang, & Liu, 2019). Despite the differences between these Machine Learning variations, both can be enhanced by ensemble learning (Y. Li, Wang, & Xu, 2018) . This research is on integrating some aspect ensemble learning in context of Deep Learning against traditional Machine Learning to unveil any possible contrasting insights.

A possible approach to ensemble learning is to integrate existing standard algorithms. For example:(1) CNN was combined with RNN to predict better crop yield prediction(Khaki, Wang, & Archontoulis, 2020); (2) better prediction of soya bean yield was observed when CNN was integrated with LSTM (Sun, Di, Sun, Shen, & Lai, 2019); and (3) Adaboost generated improved text analysis with datasets such as Reuters-21578 (Bloehdorn & Hotho, 2006). This approach however has not always guaranteed performance improvement (W. Li et al., 2019).

Another approach towards ensemble learning is to introduce an innovation and blend it with existing algorithms. For example: (1) ontology was integrated with naive Bayes classifier to enhance classification(Chang & Huang, 2008), and (2)Gaussian function integrated into CNN predicted better bean yield (Sabini, Rusak, & Ross, 2017). This second approach also a suffers a similar a fate as the first; innovations introduced provide no guarantee for performance improvement. A common opportunity in both approaches is the wide continuum of un explored options though this does not interfere with researchers interested in fixing failed cases on the two approaches. This research is towards unexplored option in the second approach.

An innovation brought forth imitates human cognition. In a text classification task, humans mentally ignore certain text and give more attention to others depending whether the portion of text read is syntactically or semantically contributing to a decision if the entire text belongs to some class. This allows humans to use less mental energy and to categorise better a text to some class of interest. Human acquire this skill through personal experience. This is plausible reason why humans out smart computers in text classification. The innovation was modeled and appropriately fitted into a text classification process. The process has phases such as; pre-processing, classifier training, classifier selection and evaluation. It is at the pre-processing stage that the innovation is introduced. Pre-processing deals with; cleaning, removal of irrelevant text, and formatting text for algorithmic manipulation. The innovation is on elimination less relevant text. Previous studies have also worked towards address the knowledge gap of irrelevant text removal. For example: Makrehchi and Kamel (2017) proposed backward filter-level as a sophisticated technique to spot irrelevant words to enhance performance, and (2) Baradad and Mugabushaka (2015) showed that using Poisson distribution less irrelevant text can also be eliminated to enhance performance. In this study similarity algorithms are manipulated to select more relevant text as describe in methodology section.

News articles were used to explore the efficacy of the innovation. These articles were containing food as keyword and were scraped from a news site and were for years 2015 and 2016. Articles containing the selected keyword word can potentially be on food insecurity or otherwise. Food insecurity is a pertinent situation to investigate as it is a big challenge afflicting more than 820 million people in world and has also attracted the attention of several researcher(FAO, IFAD, UNICEF, WFP, & WHO, 2019; Guma, Rwashana, & Oyo, 2017). To make study manageable Uganda was used as case study, a nation which is also not immune to food evidenced by several studies(for example see Dietz, 1987; Food and Agriculture Organisation, 2002; IPC, 2017; Okori et al., 2009; UmanaAponte, 2011).

The following contributions are brought forth:

1. In the context of using traditional ML or Deep Learning to study food insecurity/food dynamics, this study closely relates to several studies (Lukyamuzi, Ngubiri, & Okori, 2018, 2015; Okori, Obua, & Quinn, 2011; Quinn, Okori, & Gidudu, 2010; Surjandari, Naffisah, & Prawiradinata, 2015). These studies utilized ; tweets, phone data, satellite imagery, and food prices. This study is on utilizing news articles to study food insecurity using the unexplored blended approach.
2. The study relates further to studies that have introduced innovations to enhance performance of existing algorithms (J. Chen, Hu, Liu, Xiao, & Jiang, 2019; M. Liu & Yang, 2012; Narayanan, Arora, & Bhatia, 2013). This study proposes a technique to imitate human cognitive abilities so as to enhance performance as described.
3. Lastly the study related to studies that have also explored comparisons in algorithmic performance between traditional Machine Learning and Deep Learning (Baroni, Dinu, & Kruszewski, 2014; Hassan & Mahmood, 2017; Paterakis, Mocanu, Gibescu, Stappers, & Alst, 2018). This study also pursues a similar venture though it integrates some aspect of ensemble learning which not many studies have done.

The rest of this paper is sequentially organized as: related work; methodology; results & discussion; and conclusion & future work.

RELATED WORK

Traditional Machine Learning

Text manipulations require conversion of text into some numerical format before subjecting them to algorithms (Mandelbaum & Shalev, 2016). Numerical representations selected for texts plays a key role in text mining performance and therefore this is worthwhile issue to study. Hot vector approach is a traditional method for representing text in numerical vector form. It can be based on Term Frequency (TF) or Term Frequency Inverse Document Frequency (TF-IDF) operating on Bag-Of-Words (BOWS). TF based counts lack normalization and this causes exaggeration effect of high frequency words. TF-IDF is commonly preferred as it attempts to control the effect of high frequency words. Bag-Of-Words (BOWS) model ignores word order and this is a limiting factor. Using n-gram with n greater than 1 attempt to control loss of word order however this increases vocabulary volume and vector sparsity which calls for high computation resources.

Use of hot vector representations in the context of BOWs has yielded successful results in text classification tasks. Trstenjak et al.(2014) generated promising findings with TF-IDF and KNN to classify documents into sports, politics, finance and daily news. However in this research word semantics was ignored which sabotage classification task. In the research by Chang & Huang (2008) semantic of words is exploited using ontology. Findings revealed enhanced performance. However an ontology can be sophisticated to use; for example it requires knowledge on word senses such as polysemy, synonymy and part-of-speech tagging (Islam & Inkpen, 2008; Khan, Baharudin, Lee, & Khan, 2010). In addition, different domains may require appropriate ontology whose guarantee about their availability is not certain. Word negation has been known to cause text classification challenges. In a research by Narayanan et al., (2013) accuracy improvement in Naïve Bayes classifier is generated with strategy to handle word negation situation. Liu & Yang (2012) modified TF-IDF to TF-IDF-CF, where CF stands for class frequency. This resulted into better classification accuracy with datasets; Reuters-21578 and 20newsgroup. Coming back to the current study, the researchers seek to introduce a technique imitating human cognition to enhance accuracy and time complexity on TF-IDF environment. This is blended with traditional Machine Learning algorithms (such as KNN, NBAYES and D-TREES) to enhance performance in the context of predicting food insecurity. This is however not to claim that the current study is the first to focus on food insecurity. In the context of using traditional ML to study food insecurity/food dynamics, the current study closely relates to several studies (Lukyamuzi et al., 2018, 2015; Okori et al., 2011; Quinn et al., 2010; Surjandari et al., 2015). These studies propose utilization of datasets; tweets, phone data, satellite imagery, and food prices. The current study is on utilizing news articles to study food insecurity using the innovation described and this difference is bound to introduce new insights to research community

Deep Learning

Text representations into numerical format equally play an important role in deep Learning as stated in traditional ML. Word embedding --text representation in Deep Learning-- integrates both semantic and syntactic similarity among words. It is based on assumption that word meaning is defined by the company of word it keeps (Chapman & Christopher, 2005). It has generated amazing findings in tasks such as concept categorization, synonym detection, and analogy processing (Baroni et al., 2014). Dominant techniques to generate word embedding include; Glove, fastText, and Word2Vec. Glove exploits statistical information based on nonzero elements in a word-word co-occurrence matrix, and it has outperformed other models on tasks such as analogy and word similarity (Jeffrey, Socher, & Christopher, 2014). FastText breaks each word into sub-words. For example apple can be split into app, ppl and ple. This allows fastText to handle complicated and rare inflections (Bojanowski, Grave, Joulin, & Mikolov, 2017). Word2Vect was proposed by Mikolov, Chen, Corrado, & Dean (2013) and it is in two flavors; Skip-gram and Continuous Bag Of Words (CBOW). CBOW strategy builds word embedding by attempting to predict center word given a group of surrounding words. In Skim gram, instead prediction is on surrounding words given a center word. Word2vec have generated promising results with semantic and syntactic relations among words (Mikolov et al., 2013).

Now coming to word embedding in multiword text (eg sentence, phase and document), studies suggest generating text embedding using averages, sums and products of constituent word embedding in multiword text (Wang et al., 2018). Text embeddings at multiword level have generated promising findings. Jin, Yue, Chen, & Xia, (2016) introduced improvement into word-of-embedding and document-embedding to formulate Bag-of-Embedding. This was integration of Bag-of-Embedding with class information. This generated improvement in text classification when experimented with news twenty newsgroup (20NG). Major, Surkis, & Aphinyanaphongs (2018) compare classification performance between of-the-shell trained word embedding and custom generated word embedding in medical field. Their findings did not show superior performance from custom generated word embedding as their expectation. They believed a more customized corpus would generate improvement. Next, the study reviewed text classification in Deep Learning.

Now much focus was on two Deep Learning algorithms explored in this study; Convolutional Neural Networks (CNN) and Long Short Term Memory (LSTM). Convolution Neural Networks were inspired by visual system of animals. Its power came into focus when it won image classification competition in 2012 (Krizhevsky, Sutskever, & Hinton, 2012) . Since then the power of CNN has been growing (Bui, Lech, Cheng, Neville, & Burnett, 2017; S. Liu & Deng, 2015). The power of CNN has been extended to text classification. For example Mandelbaum & Shalev (2016) trained Convolutional Neural Networks on top of pre-trained word vectors. Promising findings were obtained compared with state-of-art practice on sentence-level classification tasks. Lenc & Král (2017a), showed that Convolutional Neural Networks generated superior performance compared with multi-layer perceptron on multi-label classification task of Czech newspaper corpus. These researchers further trained Convolutional Neural Networks in text classification tasks both in Czech Cˇ TK and English Reuters-21578 standard corpora. The Czech language was chosen as a representative of highly inflectional Slavic language with a free word order. These were longer texts yet previous studies had used shorter texts. Notable findings were obtained though it was observed that random initialization does not play an important role for classification compared to static approach.

To address challenges of CNN, Recurrent Convolutional Neural Networks were proposed. RNN is capable of harnessing signal due to sequential nature in some datasets such as sentences, phrases and time series data. Capturing such information is known to cause improvement in tasks such as classification, translation, and language modeling. Ordinary Neural Network have limitations in capturing such information and so are the CNNs. Lai, Xu, Liu, & Zhao, (2015) used recurrent structure of RNN to capture contextual information in learning word representations to reduce noise in text classification. At the same time they anticipated that max-pooling layer from convolution networks would automatically learn key word players in text classification. Their experimentation with four Chinese and English classification tasks generated results which outperformed Convolutional Neural Networks on all the datasets.

Long Short Term Memory (LSTM) is an enhancement on Recurrent Neural Networks (RNN). LSTM (unlike ordinary RNN) attempt to control challenges of exploding and varnishing gradient. These are challenges in tracking long range dependences in sequential and temporal datasets. Promising findings were obtained with LSTM in text classification task in mapping a questions into its specific SD REPORT (Sharma & Hazra, 2018). Li et al., (2018) integrated LSTM, CNN, and Bidirection ability to classify both English and Chinese texts using data sets including; (1) Stanford Sentiment Treebank benchmark, (2) subscription channel of Sina News RSS from 2005 to 2011, and (3) BBC news website. The integration generated better results than CNN, B-LSTM, SVM and RNN. Now coming back to this study, experimentation is on integrating LSTM with similarity algorithm to promote performance.

In the context of Deep Learning and studying food insecurity/food dynamics, the current study closely relates to some studies (Ganguli, Dunnmon, & Hau, 2019; Sabini et al., 2017; You, Li, Low, Lobell, & Ermon, 2017). These studies have used crop satellite imagery to predict crop yield which in turn can provide a clue on a pending food insecurity. The difference in this study is the use of news articles to study the same phenomena but at same time this done in the context of integrating CNN and LSTM with the proposed innovation.

Text Similarity

In the introduction it was highlighted that similarity algorithms was deployed as carrier to the proposed innovation. This is a review of text similarity measures as a background on similarity measures adopted in the study. Cosine measure can establish similarity between texts such as phrases, sentences and paragraphs. Other metrics include; Jaccard distance, Euclidean distance and Dice's coefficient (Hajeer, 2012). Cosine measure in traditional Machine Learning establishes word overlap between texts and it is based on TF-IDF. It generates scores from 1 to 0 with 1 as the highest and 0 the lowest. Unfortunately in traditional Machine Learning, cosine measure ignores word meaning. Words with similar meaning but with different spelling are treated as different. External dictionary such as WordNet can address this. For example in a study by Rahutomo, Kitasuka, & Aritsugi (2012) cosine measure was integrated with WordNet to compute text similarity and this resulted into positive findings. This current study uses cosine measure to sieve relevant text to promote classification performance in the context of traditional Machine Learning.

In deep learning environment to text similarity measure are based on word embedding which integrate both semantic and syntactic word relations. This study adopts a Word Mover's Distance(WMD) which utilizes graphical techniques between vector representations of text (Kusner, Sun, Kolkin, & Weinberger, 2015). WMD measures minimum distance of embedded words of a document to travel to reach embedded words of another document (Kusner et al., 2015). It calculates the total distance between words that are semantically or syntactically related. Other distance measures have been proposed such as word vector-based Dynamic Time Warping (wDTW) and word vector-based Tree Edit Distance (wTED) (Zhu, Klabjan, & Bless, 2017). This research adopts Word Movers' Distance to sieve relevant text in context of Deep Learning to promote classification accuracy.

METHODOLOGY

This section presents a strategy used to accomplish the study. On the basis of related literature it can be inferred that a text classification pipeline/process is composed of (1) text acquisition, (2) feature extraction, (3) classifier training, and (4) classifier evaluation (Kowsari et al., 2019; Surjandari et al., 2015). Our study required exploring use of ensemble learning while contrasting traditional Machine Learning against Deep Learning in a context of classifying if a news article is on food insecurity or not. From this the best classifier was applied with additional manipulation to generate trends of food insecurity. To match the study specifics, the text classification process was customized according. The process adjusted and it included; (1) text acquisition, (2) preprocessing, (3)training baseline classifiers, (4)training ensemble classifiers, (5)evaluating ensemble classifiers, (6) generating food insecurity trends, and (7) validation with ground truth.

Data Acquisition

The articles used were scrapped from Monitor news site (http://www.monitor.co.ug) for years 2015 and 2016. Monitor is a popular news media in Uganda and it was chosen as representative source of news articles. Articles scrapped were based on 'food' as a search keyword. Articles containing this search keyword can be on food insecurity or not. Other keywords such as crops and farmers could serve a similar purpose. The study chose to explore with food as a search keyword. For each article, the scrapper captured: article's title, posting date and the content. A total of 2664 articles were retrieved and saved in csv format. Table 1 summarizes capacities of articles retrieved.

Table 1. Showing number of articles retrieved

Year	Total Articles Retrieved
2015	1323
2016	1341
Total	2664

Preprocessing

At this stage articles were; further cleaned, and tokenized. In traditional Machine Learning, text was tokenized using in build nltk library in python. The tokens were then converted into numerical representation based on TF-IDF ready for classifier training. In Deep Learning, the tokens were generated into word embedding. Several libraries such as gensism, keras, and torch in python generate word embedding. The study used keras library to generate word embedding.

Training Baseline Classifiers

Baseline classifiers were trained on articles for 2016. Articles for 2015 were left for experimentation with best model. In traditional Machine Learning, data sets had been transformed into TF-IDF vector representations. These were subjected to four algorithms; K-Nearest Neighbors, Naïve Bayes, Support Vector Machine, and Decision Trees. Each algorithm was trained on 10 cross validation strategy. In Deep Learning, data sets were transformed into word embedding. The word embedding were then subjected to two algorithms; Convolution Neural Networks (CNN), and Long Short Term Memory (LSTM). At this stage, datasets were split into 80% for training and 20% for testing. For all these cases (traditional or deep learning), accuracies and execution times for each algorithm were computed for evaluation stage.

Training Ensemble Classifiers

Articles for 2016 were also used for classifier training and articles for 2015 were also left for experimentation with best model. Before explaining how this training was done it is important to describe the structure of these ensemble classifiers. In this study ensemble classifiers were formed by integrating text similarity algorithms with: first with traditional Machine Learning algorithms and then with Deep Learning algorithms. In brief, ensemble algorithms were structured in serial forms. An ensemble in traditional environment; preprocessed data was first subjected to similarity algorithm then the outcome was subjected to traditional classifier. An ensemble in Deep Learning; preprocessed data was first subjected to similarity algorithm then the outcome was subjected to Deep classifier.

A similarity algorithm used in traditional Machine Learning is different to that used in deep learning environment due to technological capabilities available. In traditional Machine Learning, similarity algorithm used was based on cosine measure. In Deep Learning the study employed similarity algorithm based on Word Mover's Distance (WMD). In both case similarity algorithms were used to select most relevant text for class of interest from each article; in this case focus was on conversations on food insecurity.

To select relevant text, the study required suitable text on food insecurity. Three options were available to generate the suitable text; (1) use of domain knowledge, (2) use of language modeling techniques, and (3) search and concatenating relevant sentences/phrases on food insecurity from retrieved articles. This study used the last option. Ten sentences/phrases on food insecurity were randomly selected from retrieved articles to form the suitable text. Each article was split into separate sentences and each sentence was subjected to similarity measure with the suitable text. A relevant sentence was expected to generate high similarity value. In traditional Machine Learning, cosine measure was used to generate similarities scores between suitable text and each sentence. Top four most similar sentences were retained from each article. This means each article was down sampled to four sentences. In deep Learning, WMD was instead used to determine and down sample each article to four most similar sentences.

In traditional Machine Learning, the down sampled text -- generated by cosine measure -- was subjected to the four classifiers; K-Nearest Neighbors, Naïve Bayes, Support Vector Machine, and Decision Trees. They were trained on 10 cross validation strategy. In deep learning, the down sampled text -- generated with WMD-- was on other hand subjected to the two deep learning algorithms; CNN and LSTM. In Deep Learning the training the datasets were split into 80% for training and 20% for testing. For each training (ensemble in traditional or Deep Learning) accuracy and execution time was computed.

Evaluating Ensemble Classifiers

Evaluation was based on two evaluation parameters; accuracy and execution time. Accuracies and execution times in baseline classifiers were compared to measurements with ensemble classifiers in traditional Machine Learning and Deep Learning. Higher accuracies when ensemble classifiers used would signify improvement. Reduced execution time with ensemble classifiers would on other hand signify improved speed of execution.

Generating Food Insecurity Trends

A best performing classifier was employed to label datasets for year 2015. Monthly counts for incidences of articles talking about food insecurity were computed and generated into a time series.

Validating Trends With Ground Truth

Food prices released by (Uganda Bureau of Statistics, 2015) were used to establish the ground truth in comparison with food insecurity trends. The assumption is that food price changes should be positively correlated with conversations trends on food insecurity. Due to differences in value ranges (insecurity counts and food price values) the two data sets were normalized to create some uniformity in graphic visualization. Saranya & Manikandan (2013) offers three techniques of normalizing: Z-score, Decimal scaling normalization, and Min-Max. In this study Min-Max was adopted because it is a linear and a simple technique to implement. Using this technique, data sets were fitted into pre-defined boundary. In this study, [0, 1] boundary was chosen.

RESULTS AND DISCUSSION

Results

Classifier Accuracies

The accuracies presented are both with and without innovation. A column has been added to give performance improvement with the innovation. Improvement has been computed as difference in accuracy with and without the innovation. This is detailed in the table 2. Note that the best performance was with CNN. Out of six algorithms, four of them generated performance improvement while for the rest the performance degraded.

Table 2. Showing accuracies operating at random classifier of 0.744

Classifier	Accuracy Without Innovation	Accuracy With Innovation	Accuracy Improvement
1. KNN	0.727	0.752	0.025
2. NBAYES	0.772	0.786	0.014
3. SVM	0.7437	0.747	0.003
4. D-TREES	0.758	0.788	0.030
5. CNN	0.804	0.79	-0.005
6. LSTM	0.780	0.750	-0.03

Computation Times

All algorithms resulted into reduction in computation times by a factor 2.9 and above. LSTM resulted into the highest reduction in time by a factor of 22. All this is observable in table 3.

Validating Trends With Alternative Sources

Trends of food insecurity from news articles were compared against food price trends from Uganda Bureau of Statistics. Comparison was with two cases. First trends due to human annotation against price trends as shown by Figure 1. The correlation here was 0.473. The second case was the correlation due automatic annotation using CNN (best classifier) against the same price trends as shown in Figure 2. The correlation for these was 0.449.

Discussion

It was found that the proposed innovation generated improvement in accuracy. Four algorithms out of six portrayed performance improvement (See table 2). This relates to researches such as (Hassan & Mahmood, 2017; Khedr, Kadry, & Walid, 2015; Y. Li et al., 2018; Narayanan et al., 2013). These investigations introduced innovations to improve accuracy and in some cases the computation times were reduced. This study used a different innovation; it was based on imitating human cognition to promote accuracy. Similarity algorithms were re-used to achieve this. They guided classifiers to

Table 3. Computation times with innovation and without innovation

Classifier	Runtime with No-Innovation	Runtime with Innovation	Runtime Ration: with Innovation to No-Innovation
1. KNN	0.945	0.292	3.2
2. NBAYES	0.156	0.066	2.3
3. SVM	0.036	0.009	4.0
4. D-TREES	0.012	0.004	2.8
5. CNN	72	77	7.2
6. LSTM	1149	129	8.9

Figure 1. Trends of food insecurity incidences; manually labeled against UBOS food prices

Figure 2. Trends of food insecurity incidences; automatically labeled against UBOS food prices

focus on relevant text. The innovation was explored both in traditional Machine Learning and in deep learning setting. To the best of reseachers' knowledge they have not come across a research which has explored this possibility moreover in the context of tracking food insecurity using news articles. More so, in this context the study integrated some aspect of ensemble learning; each of these six algorithms was blended with appropriate similarity algorithm. The proposed technique can be used in conjunction with previous innovation(s) to put performance to higher level. A relevant question to ask is: How significant is the improvement? This is a tough question as the performance improvement was not very exciting. More so the best performing model was without the innovation. CNN without the innovation generated the best accuracy. This is the reasons why this CNN was deployed in generating trends of food insecurity. At least the overall picture of the proposed ensemble is projecting some light (portrayed by performance improvement) at the end of the tunnel. It is worthwhile to shift the attention on how to brighten this light. Pursuing this, calls for further investigations as suggested in last section.

The two deep learning algorithms explored portrayed retardation in performance with the innovation. A plausible reason is due to their sophistication. Deep learning algorithms are advanced in teasing patterns but this is mainly possible in data intensive environment. Reduction in data due to innovation possibly turned out to be a limiting factor. Nevertheless, retardation does not render the innovation completely useless. Tradeoffs can be made. For example, in Netflix competition the best performing was not adopted by the company. Instead the 2nd best algorithm with lower accuracy was adopted after making a trade off because this algorithm demanded favorable resources. Does a similar situation apply to this research? The innovation is in its infancy, possibly more experimentation still deserves some consideration. All six algorithms portrayed faster execution by a factor at least greater than two (see table 3). Time saving ranged from seconds to minutes. There was no time saving close to an hour. This is because the datasets used were relative smaller compared to industrial scale experimentation. On industrial scale, execution time can go on for days. This is illustrated in a study by Mikolov et al., (2013). Now assuming a linear reduction in time complex, this mean the proposed innovation would reduce execution period of two days to at least day. This is a worthwhile time saving.

Visualizing these food insecurity trends from news articles were comparable to food price trends from UBOS as an official source. Food prices were chosen because they are capable of providing proxy information about food insecurity. The price trends were first compared with food insecurity trends due to human annotation and this was at correlation of 0.473 (see Figure 1). Secondly, food price trends were compared with food insecurity trends due to automatic annotation and this was at correlation of 0.449 (see Figure 2). The two correlation coefficient results are comparable with previous studies (Lukyamuzi et al., 2018; Malhotra & Maloo, 2017; Pulse UN Global, 2014). In (Pulse UN Global, 2014), using tweets on food price trends generated a correlation of 0.42 compared to official datasets. In a research by Lukyamuzi et al.(2018), using tweets conversations on food insecurity generated a correlation of 0.56 for year 2015 and 0.37 for year 2016 compared to food price inflation from official source, UBOS. What is the possibility of improving these correlations? Possible opportunities are highlighted in future work.

CONCLUSION AND FUTURE DIRECTION

The research set out to investigate if human cognitive abilities is blended with classifiers would improve accuracy and reduce computation time. To imitate this cognitive ability, appropriate similarity algorithms, were blended with six classifiers . The innovation was experimented in both traditional Machine and Deep Learning in context of tracking food insecurity. Promising findings have been obtained though not very exciting as the best model was without innovation. Consequently CNN (the best model) was deployed in tracking food insecurity. Performance was comparable with ground truth of tracking food insecurity using food prices. Two suggestions are put forth to promote performance: (1) using articles aggregated from several news media, and (2) blending more classifiers in an ensemble.

REFERENCES

Abdel Aziz, M. S. E.-D. A., Elsamahy, M., Hassan, M. A. M., & Bendary, F. M. A. (2016). Enhancement of Turbo-Generators Phase Backup Protection Using Adaptive Neuro Fuzzy Inference System. *International Journal of System Dynamics Applications*, *6*(1), 58–76. doi:10.4018/IJSDA.2017010104

Andreas, T., Jahrer, M., Bell, R. M., & Park, F. (2009). *The BigChaos Solution to the Netflix Grand Prize*. Retrieved from https://www.netflixprize.com/assets/GrandPrize2009_BPC_BigChaos.pdf

Baradad, V. P., & Mugabushaka, A. (2015). Corpus Specific Stop Words to Improve the Textual Analysis in Scientometrics. *15th International Conference of the International Society for Scientometrics and Informetrics (ISSI)*, 999–1005. Retrieved from https://pdfs.semanticscholar.org/618d/a4f6d5cd329d3bc498d9457f575cbdfaf53d.pdf

Baroni, M., Dinu, G., & Kruszewski, G. (2014). Don't count, predict! A systematic comparison of context-counting vs. context-predicting semantic vectors. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, 238–247. doi:10.3115/v1/P14-1023

Bloehdorn, S., & Hotho, A. (2006). *Boosting for Text Classification with Semantic Features*. Retrieved from https://link.springer.com/chapter/10.1007/11899402_10

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching Word Vectors with Subword Information. *Transactions of the Association for Computational Linguistics*, *5*, 135–146. doi:10.1162/tacl_a_00051

Bouzaida, S., & Sakly, A. (2018). Adaptive Neuro-Fuzzy Sliding Mode Controller. *International Journal of System Dynamics Applications*, *7*(2), 34–54. doi:10.4018/IJSDA.2018040103

Bui, H. M., Lech, M., Cheng, E., Neville, K., & Burnett, I. S. (2017). Object Recognition Using Deep Convolutional Features Transformed by a Recursive Network Structure. *IEEE Access: Practical Innovations, Open Solutions*, *4*, 10059–10066. doi:10.1109/ACCESS.2016.2639543

Chang, Y.-H., & Huang, H.-Y. (2008). An Automatic Document Classifier System based on Naíve Bayes Classifier and Ontology. *Proceedings of the 7th International Conference on Machine Learning and Cybernetics, ICMLC*.

Chapman, S., & Christopher, R. (2005). Firth, J.R. (John Rupert). *Key Thinkers in Linguistics and the Philosophy of Language*, (4), 80–86. Retrieved from http://www.lel.ed.ac.uk/homes/patrick/firth.pdf

Chen, J., Hu, Y., Liu, J., Xiao, Y., & Jiang, H. (2019). *Deep Short Text Classification with Knowledge Powered Attention*. Academic Press.

Chen, Q., & Sokolova, M. (2018). *Word2Vec and Doc2Vec in Unsupervised Sentiment Analysis of Clinical Discharge Summaries*. Retrieved from https://arxiv.org/abs/1805.00352

Dietz, A. J. (1987). *Pastoralists in dire straits : survival strategies and external interventions in a semi-arid region at the Kenya/Uganda border: Western Pokot, 1900-1986*. Retrieved from https://openaccess.leidenuniv.nl/handle/1887/68673

FAO. (2019). *IFAD, UNICEF, WFP, & WHO*. Food Security and Nutrition in the World., doi:10.1109/JSTARS.2014.2300145

Feuerverger, A., He, Y., & Khatri, S. (2012). Statistical Significance of the Netflix Challenge. *Statistical Science*, *27*(2), 202–231. doi:10.1214/11-STS368

Food and Agriculture Organisation. (2002). *The State of Food Insecurity in the World*. Retrieved from http://www.fao.org/3/y7352e/y7352e03.htm#TopOfPage

Ganguli, S., Dunnmon, J., & Hau, D. (2019). *Predicting Food Security Outcomes Using CNNs for Satellite Tasking*. Retrieved from https://arxiv.org/pdf/1902.05433.pdf

Guma, I. P., Rwashana, A. S., & Oyo, B. (2017). Food Security Indicators for Subsistence Farmers Sustainability. *International Journal of System Dynamics Applications*, *7*(1), 45–64. doi:10.4018/IJSDA.2018010103

Hajeer, S. I. (2012). Comparison on the Effectiveness of Different Statistical Similarity Measures. *International Journal of Computers and Applications*, *53*(8), 14–19. doi:10.5120/8440-2224

Hallinan, B., & Striphas, T. (2016). Recommended for you : The Netflix Prize and the production of algorithmic culture. *New Media & Society*, *18*(1), 117–137. doi:10.1177/1461444814538646

Hassan, A., & Mahmood, A. (2017). Deep Learning approach for sentiment analysis of short texts. *2017 3rd International Conference on Control, Automation and Robotics Deep*. doi:<ALIGNMENT.qj></ALIGNMENT>10.1109/ICCAR.2017.7942788

He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition*. <ALIGNMENT.qj></ALIGNMENT>10.1016/0141-0229(95)00188-3

Hussein, H. A. T., Ammar, M. E., & Hassan, M. A. M. (2017). Three Phase Induction Motor's Stator Turns Fault Analysis Based on Artificial Intelligence. *International Journal of System Dynamics Applications*, *6*(3), 1–19. doi:10.4018/IJSDA.2017070101

IPC. (2017). *Current Acute Food Insecurity Situation*. IPC.

Islam, A., & Inkpen, D. (2008). Semantic Text Similarity Using Corpus-Based Word Similarity and String Similarity. *ACM Transactions on Knowledge Discovery from Data*, *2*(2), 1–25. Advance online publication. doi:10.1145/1376815.1376819

Jeffrey, P., Socher, R., & Christopher, D. M. (2014). GloVe: Global Vectors for Word Representation. *Proceedings Ofthe 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543. doi:<ALIGNMENT.qj></ALIGNMENT>10.3115/v1/D14-1162

Jin, P., Yue, Z., Chen, X., & Xia, Y. (2016). Bag-of-embeddings for text classification. *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16),* 2824–2830. doi:<ALIGNMENT.qj></ALIGNMENT>10.1080/07421222.2014.995538

Khaki, S., Wang, L., & Archontoulis, S. V. (2020). A CNN-RNN Framework for Crop Yield Prediction. *Frontiers in Plant Science*, *10*(2), 1750. Advance online publication. doi:10.3389/fpls.2019.01750 PMID:32038699

Khan, A., Baharudin, B., Lee, L. H., & Khan, K. (2010). A Review of Machine Learning Algorithms for Text-Documents Classification. *Journal of Advances in Information Technology*, *1*(1), 4–20. doi:10.4304/jait.1.1.4-20

Khedr, A. E., Kadry, M., & Walid, G. (2015). Proposed framework for implementing data mining techniques to enhance decisions in agriculture sector Applied case on Food Security Information Center Ministry of Agriculture, Egypt. *Procedia Computer Science*, *65*, 633–642. doi:10.1016/j.procs.2015.09.007

Kowsari, K., Meimandi, K. J., Heidarysafa, M., Mendu, S., Barnes, L., & Brown, D. (2019). Text Classification Algorithms : A Survey. *Information*, *10*(4), 1–68. doi:10.3390/info10040150

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *ImageNet Classification with Deep Convolutional Neural Networks*. Retrieved from https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

Kusner, M. J., Sun, Y., Kolkin, N. I., & Weinberger, K. Q. (2015). From Word Embeddings To Document Distances. *Proceedings of The 32nd International Conference on Machine Learning, 37*, 957–966. Retrieved from http://proceedings.mlr.press/v37/kusnerb15.pdf

Lai, S., Xu, L., Liu, K., & Zhao, J. (2015). Recurrent Convolutional Neural Networks for Text Classification. *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence Recurrent*, 2267–2273. Retrieved from https://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9745/9552

Lenc, L., & Král, P. (2017). Deep Neural Networks for Czech Multi-label Document Classificatio. *CEUR Workshop Proceedings*. doi:<ALIGNMENT.qj></ALIGNMENT>10.1007/978-3-319-59569-6_34

Li, W., Liu, P., Zhang, Q., & Liu, W. (2019). An Improved Approach for Text Sentiment Classification Based on a Deep Neural Network via a Sentiment Attention Mechanism. *Future Internet*, *11*(4), 96. doi:10.3390/fi11040096

Li, Y., Wang, X., & Xu, P. (2018). Chinese Text Classification Model Based on Deep Learning. *Future Internet*, *10*(11), 113. Advance online publication. doi:10.3390/fi10110113

Liu, M., & Yang, J. (2012). An improvement of TFIDF weighting in text categorization. *International Conference on Computer Technology and Science (ICCTS 2012)., 47*(2012), 44–47. doi:<ALIGNMENT.qj></ALIGNMENT>10.7763/IPCSIT.2012.V47.9

Liu, S., & Deng, W. (2015). Very Deep Convolutional Neural Network Based Image Classification Using Small Training Sample Size. *2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, 7486599. Retrieved from https://ieeexplore.ieee.org/document/7486599/authors#authors

Lukyamuzi, A., Ngubiri, J., & Okori, G. (2018). Tracking Food Insecurity from Tweets Using Data Mining Techniques. *SEIA '18: SEIA '18: Symposium on Software Engineering in Africa*. doi:10.1145/3195528.3195531

Lukyamuzi, A., Ngubiri, J., & Okori, W. (2015). Towards Harnessing Phone Messages and Telephone Conversations for Prediction of Food Crisis. *International Journal of System Dynamics Applications*, *4*(4), 1–16. doi:10.4018/IJSDA.2015100101

Majhi, S. K. (2018). An Efficient Feed Foreword Network Model with Sine Cosine Algorithm for Breast Cancer Classification. *International Journal of System Dynamics Applications*, *7*(2), 1–14. doi:10.4018/IJSDA.2018040101

Major, V., Surkis, A., & Aphinyanaphongs, Y. (2018). Utility of General and Specific Word Embeddings for Classifying Translational Stages of Research. *AMIA 2018 Annual Symposium.* Retrieved from https://arxiv.org/ftp/arxiv/papers/1705/1705.06262.pdf

Makrehchi, M., & Kamel, M. S. (2017). Extracting domain-specific stopwords for text classifiers. *Intelligent Data Analysis*, *21*(1), 39–62. doi:10.3233/IDA-150390

Malhotra, A., & Maloo, M. (2017). Understanding Food Inflation in India: A Machine Learning Approach. *SSRN Electronic Journal*, 1–44. <ALIGNMENT.qj></ALIGNMENT>10.2139/ssrn.2908354

Mandelbaum, A., & Shalev, A. (2016). *Word Embeddings and Their Use In Sentence Classification Tasks*. Academic Press.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient Estimation of Word Representations in Vector Space*. <ALIGNMENT.qj></ALIGNMENT>10.1162/153244303322533223

Narayanan, V., Arora, I., & Bhatia, A. (2013). Fast and accurate sentiment classification using an enhanced Naive Bayes model. Lecture Notes in Computer Science, 8206, 194–201. doi:<ALIGNMENT.qj></ALIGNMENT>10.1007/978-3-642-41278-3_24

NETFLIX. (2009). *Leaderboard*. Retrieved from https://www.netflixprize.com/leaderboard.html

Okori, W., Obua, J., & Baryamureeba, V. (2009). Famine Disaster Causes and Management Based on Local Community ' s Perception in Northern Uganda. *Research Journal of Social Sciences*, *4*, 21–32.

Okori, W., Obua, J., & Quinn, J. (2011). Machine Learning Classification Technique for Famine Prediction. *Proceedings of the World Congress on Engineering, 2*(1), 4–9.

Paterakis, N. G., Mocanu, E., Gibescu, M., Stappers, B., & van Alst, W. (2018). Deep learning versus traditional machine learning methods for aggregated energy demand prediction. *Computer and Information Sciences (ICCOINS) 2018 4th International Conference On*, 1–6.

Pulse U. N. Global. (2014). *Mining Indonesian Tweets to Understand Food Price Crises*. Retrieved from http://www.unglobalpulse.org/sites/default/files/Global-Pulse-Mining-Indonesian-Tweets-Food-Price-Crises copy.pdf

Quinn, J. A., Okori, W., & Gidudu, A. (2010). Increased-specificity famine prediction using satellite observation data. *Proceedings of the 1st ACM Symposium on Computing for Development, DEV 2010*. doi:10.1145/1926180.1926203

Rahutomo, F., Kitasuka, T., & Aritsugi, M. (2012). Semantic Cosine Similarity. *The 7th International Student Conference on Advanced Science and Technology ICAST*.

Ren, Y., Zhang, L., & Suganthan, P. N. (2016). Ensemble Classification and Regression-Recent Developments, Applications and Future Directions [Review Article]. *IEEE Computational Intelligence Magazine*, *11*(1), 41–53. doi:10.1109/MCI.2015.2471235

Sabini, M., Rusak, G., & Ross, B. (2017). *Understanding Satellite-Imagery-Based Crop Yield Predictions*. Academic Press.

Saranya, C., & Manikandan, G. (2013). A study on normalization techniques for privacy preserving data mining. *IACSIT International Journal of Engineering and Technology*, *5*(3), 2701–2704.

Sharma, A. K., & Hazra, S. (2018). Application of Deep Learning Techniques for Text Classification on Small Datasets. *International Journal of Engineering Science and Computing*, *8*(4), 17212–17213.

Sun, J., Di, L., Sun, Z., Shen, Y., & Lai, Z. (2019). County-Level Soybean Yield Prediction Using Deep CNN-LSTM Model. *Sensors (Basel)*, *19*(20), 1–21. doi:10.3390/s19204363 PMID:31600963

Surjandari, I., Naffisah, M. S., & Prawiradinata, M. I. (2015). Text Mining of Twitter Data for Public Sentiment Analysis of Staple Foods Price Changes. *Journal of Industrial and Intelligent Information*, *3*(3), 253–257. doi:10.12720/jiii.3.3.253-257

Trstenjak, B., Mikac, S., & Donko, D. (2014). KNN with TF-IDF based framework for text categorization. *Procedia Engineering, 69*(October), 1356–1364. <ALIGNMENT.qj></ALIGNMENT>10.1016/j.proeng.2014.03.129

Uganda Bureau of Statistics. (2015). *Consumer Price Index*. Retrieved from http://www.ubos.org/onlinefiles/uploads/ubos/cpi/cpiDec2015/FINAL CPI Release December 2015.pdf

UmanaAponte, M. (2011). Long-term effects of a nutritional shock: the 1980 famine of Karamoja, Uganda. Academic Press.

Wang, G., Li, C., Wang, W., Zhang, Y., Shen, D., Zhang, X., … Carin, L. (2018). *Joint Embedding of Words and Labels for Text Classification*. Retrieved from https://arxiv.org/pdf/1805.04174.pdf

Wind, D. K., & Winther, O. (2014). Model selection in data analysis competitions. *CEUR Workshop Proceedings*, *1202*(1201), 55–60.

You, J., Li, X., Low, M., Lobell, D., & Ermon, S. (2017). Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data. *Thirty-First AAAI Conference on Artificial Intelligence*.

Zhu, X., Klabjan, D., & Bless, P. (2017). Semantic Document Distance Measures and Unsupervised Document Revision Detection. *Proceedings of the 8th International Joint Conference on Natural Language Processing*, 947–956. Retrieved from https://arxiv.org/abs/1709.01256

Washington Okori has over eighteen years of leadership experience with twelve years at Senior Management level. They have coordinated multiple partner organizations to delivery of high impact programs. Supervised transformation and service delivery projects that involved both local and international companies delivered on time, under budget with excellent results. They are a researcher who has carried out extensive research in disaster risk management, computational technologies used in prediction and management of disasters with imaginative goals to save life and property of those at risk. They have taught and mentored students at many Universities such as Makerere University, UTAMU and UMI in addition to review of educational curriculum for higher educational institutions.