

## Social License to Operator Triple-Bottom-Line Tweet Classification

The application domain is the Triple-Bottom-Line (TBL) classification of Tweet in the context of the Social License to Operate (SLO) of mining companies. The objective of this project is to continue and extend the earlier work on Tweet TBL classification done at CSIRO – the Commonwealth Scientific and Industrial Research Organization (Australia’s National Science Agency). The goal is to set up a prototype machine learning system that is capable of identifying the topic classification of a Tweet as either Environmental, Social, or Economic. The initial milestone is to achieve at an absolute minimum a 50% accuracy metric or higher, indicating the ability to at least guess on par with a flip of a coin.

The Social License to Operate is defined as when an existing project has the ongoing approval of the local community and other stakeholders within the domain the project operates in. It is the ongoing social acceptance of that project in regards to a favorable or dis-favorable disposition by those who are concerned. The SLO must not only be earned but also maintained as the beliefs, opinions, and perceptions of people tend to be dynamic over the course of time. It is beneficial to the project owners and managers to maintain an agreeable relationship with the local population and their stakeholders.

The Triple Bottom Line is defined as a framework where organizations and companies dedicate themselves not only to profit but also the social and environmental impact of their operation. The phrase was coined by the British management consultant John Elkington as a metric to measure the performance of corporate America. According to Investopedia, the corporate business should be done according to:

Profit – the traditional measure of corporate profit – the profit and loss (P & L) account.

People – the measure of how socially responsible an organization has been throughout its operations.

Planet – the measure of how environmentally responsible a firm has been.

These are the three elements of TBL which are then sourced into the terms Economy (profit), Environmental (planet), and Social (people).

Twitter data (Tweets) can be obtained in 4 distinct ways – retrieval from the Twitter public API, use of an existing Twitter dataset, purchase from Twitter directly, or purchase access from a 3rd party Twitter service provider. For the purposes of this project, we will be using existing Twitter datasets provided by Professor VanderLinden via access to Calvin College’s Borg supercomputer. Specifically, we will be using a training set consisting of crowdsourced Triple Bottom Line labeled Tweets used by CSIRO in their preliminary topic classification research. For the test set, we will be using a small dataset consisting of TBL labeled Tweets hand-labeled by Professor VanderLinden. With the machine learning model trained on these two sets, we will then generalize the model to make predictions on the dataset used for stance classification of Tweets in earlier research by Professor VanderLinden and Roy Adams.

As our research is a continuation of prior research from CSIRO and based on the foundation laid by Professor VanderLinden’s “Machine Learning for Social Media” project, we see no reason to not use machine learning. While we might consider symbolic artificial intelligence (GOFAI – Good, Old-Fashioned AI), we learned in CS-344 that symbolic reasoning implementations resulted in rules engines, also known as expert systems or knowledge graphs. These proved to be too brittle and became unmanageable as the knowledge base grew beyond a few thousand rules. Considering the nature of Tweets, GOFAI seems not to be a viable solution. The language of Tweets is often informal, prone to slang, misuse of established grammatical rules, and in general a chimeric bastardization of known human languages (insofar in my experience). It is doubtful a purely symbolic AI would be computationally feasible. Perhaps as Professor VanderLinden mentioned, a hybrid A.I. combining symbolic reasoning and deep neural networks is the future of A.I. and would prove to be a feasible approach.

Preliminary analysis of the two provided datasets indicates that they will require significant pre-processing before becoming useable as input features for machine learning. The Tweets are stored as comma delimited CSV files. The training dataset consists of 299 total Tweets, of which 198 are unlabeled due to not being associated with any TBL classification. The test dataset consists of 31 hand-labeled Tweets. Based on the size of the datasets we are working with neural networks may not be the best choice to start with. Neural networks typically require larger datasets in order to train and as we barely have 330 total examples to work with, the results may be less than optimal. Therefore, we will start with Bayesian models and SVM’s – Support Vector Machines. Later, we will expand to using supervised neural networks just to see if we can tune hyperparameters to obtain results closely comparable to our non-NN models.

For fast prototyping, we will be using Scikit-Learn in Python rather than Keras or straight Tensorflow, at least until we have established which baseline supervised learning algorithm will provide us with the potential for the best results. We will also use Pandas, built on NumPy, for data-frame manipulation and matplotlib for visualizations. To encode our categorical Tweet data into useable numerical Tweet data, we will be using the tools provided by Scikit-Learn.

Our Bayesian model will be the MultinomialNB classifier that implements the naïve Bayes algorithm for multinomially distributed data. Scikit-Learn.org indicates that it is one of the two classic Naïve Bayes variants used in textual classification problems. This indicates it will be an excellent starting point as we have decided our two datasets are too small to initially warrant the use of a supervised neural network training algorithm. “Naïve” in this case indicates the application of Bayes’ theorem with the “naïve” assumption of conditional independence between every pair of features given the value of the class variable (4). Further information indicates the classifier performs fast and works in many real-world applications, including document classification and spam filtering. We built a spam filter based on Paul Graham’s “A Plan for Spam” and indeed it worked well.

Our SVM classification model will be the LinearSVC Classifier— Linear Support Vector Classification. Sci-Kit Learn indicates it is effective in high dimensional spaces and when the number of dimensions is greater than the number of samples. This will be the case for us as we have a limited 330 samples and after multi-hot encoding to form a feature vector to create a bag-of-words vocabulary, our dimensionality is bound to be pretty high in comparison to the samples. The memory efficiency of this algorithm should also help as we will no doubt have sparse vectors in comparison to the total vocabulary present across all of the Tweets. Of note, is that SVM algorithms are not scaling invariant, so data scaling is required, which should not matter in our case as encoding our categorical word data will result in 0,1's for the input feature vector. API documentation indicates that the classifier supports sparse input (good for us) and supports multi-class using the one-vs-the-rest scheme.

Our deep neural network will be the MLP Classifier – multi-layer perceptron. Scikit-Learn indicates it uses a Softmax layer as the output function to perform multi-class classification and uses the cross-entropy loss function. MLP also supports multi-label classification through use of the logistic activation function where values  $> 0.5 \rightarrow 1$  and values  $< 0.5 \rightarrow 0$ . Given this, it would be possible for us to perform multi-class multi-label TBL classification on our training dataset. Our training dataset does possess Tweets that have been given multiple topic classifications, although some are redundant duplicates of either economic, social, or environmental. We will leave this possibility for the future, time permitting. Effective use of the MLP classifier would most likely require us to hand-label additional training example from the larger Twitter datasets present on Calvin's Borg supercomputer. Crowdsourcing does not seem a viable option so this task would be tediously time-consuming.

The application of machine learning to Social License to Operate on Triple-Bottom-Line topic classification can potentially assist any organization or company in evaluating their current level of acceptability by the local population and relevant stakeholders. Specifically, it could help evaluate whether people are more concerned about the economic, social, or environmental aspects of the project. In conjunction with stance and sentiment SLO machine learning models, it should be plausible that the level of acceptability of a project can be accurately judged.

With social media so prevalent in this day and age, it is a simple matter to obtain fresh new datasets on a daily basis to gauge the SLO. As such, the synchronicity between the dynamism of maintaining the SLO and new Tweets pertaining to the associated project works well. Rather than conduct old fashioned mail surveys, which is time-consuming and potentially expensive, the entire procedure can be automated. Extract Twitter data using the Twitter API, pre-process the dataset, post-process the dataset, insert into the machine learning model(s) as input feature vectors, and predict the level of approval. Given a good model, any organization, corporation, or other entity, can perform a pseudo-real-time estimate on how accepted their current operations and activities are.

The initial investment would be in adjusting hyperparameters with the validation set to achieve the optimal results while avoiding overfitting and ensuring the model generalizes well to new data. Once this is achieved, the model should be relevant and usable as an SLO predictor for a given period of time for a particular project and organization. Of course, even with a good model perhaps the best way to judge SLO would still be to do a face-to-face interview with the individuals in the community and stakeholders and simply ask how they feel about the project. Then again, the anonymity of the Internet does provide an outlet for people to vent and voice their opinions with less fear of reprisal than in reality. So perhaps anonymous Tweeters are more honest. But, anonymity could also cause people to simply say whatever they desire with little regard to how their words actually correlate to their own personal beliefs and opinions on the matter. Either way, an SLO TBL machine learned prediction model won't be the be all and end all in estimating Social License to Operate. But, it can be a useful cog in the whole machine in order to generate the necessary analysis required to measure the components of SLO.

### Works Referenced

- 1) Anonymous ACL submission. "Classifying Stance Using Profile Texts".
- 2) "1. Supervised Learning¶." *Scikit*, [scikit-learn.org/stable/supervised\\_learning.html#supervised-learning](https://scikit-learn.org/stable/supervised_learning.html#supervised-learning).
- 3) "A Gentle Introduction to the Bag-of-Words Model." *Machine Learning Mastery*, 12 Mar. 2019, [machinelearningmastery.com/gentle-introduction-bag-words-model/](https://machinelearningmastery.com/gentle-introduction-bag-words-model/).
- 4) "Introduction to Machine Learning | Machine Learning Crash Course | Google Developers." *Google*, Google, [developers.google.com/machine-learning/crash-course/ml-intro](https://developers.google.com/machine-learning/crash-course/ml-intro).
- 5) Kenton, Will. "How Can There Be Three Bottom Lines?" *Investopedia*, Investopedia, 9 Apr. 2019, [www.investopedia.com/terms/t/triple-bottom-line.asp](https://www.investopedia.com/terms/t/triple-bottom-line.asp).
- 6) Littman, Justin. "Where to Get Twitter Data for Academic Research." *Social Feed Manager*, 14 Sept. 2017, [gwu-libraries.github.io/sfm-ui/posts/2017-09-14-twitter-data](https://gwu-libraries.github.io/sfm-ui/posts/2017-09-14-twitter-data).
- 7) Mohammad, Saif, et al. "SemEval-2016 Task 6: Detecting Stance in Tweets." *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, 2016, doi:10.18653/v1/s16-1003.
- 8) "Multiclass Classification." *Wikipedia*, Wikimedia Foundation, 18 Apr. 2019, [en.wikipedia.org/wiki/Multiclass\\_classification](https://en.wikipedia.org/wiki/Multiclass_classification).
- 9) "Symbolic Reasoning (Symbolic AI) and Machine Learning." *SkyMind*, [skymind.ai/wiki/symbolic-reasoning](https://skymind.ai/wiki/symbolic-reasoning).
- 10) Walker, Leslie. "Learn Tweeting Slang: A Twitter Dictionary." *Lifewire*, Lifewire, 8 Nov. 2017, [www.lifewire.com/twitter-slang-and-key-terms-explained-2655399](https://www.lifewire.com/twitter-slang-and-key-terms-explained-2655399).
- 11) "What Is the Social License?" *The Social License To Operate*, [sociallicense.com/definition.html](https://sociallicense.com/definition.html).
- 12) "Working With Text Data¶." *Scikit*, [scikit-learn.org/stable/tutorial/text\\_analytics/working\\_with\\_text\\_data.html](https://scikit-learn.org/stable/tutorial/text_analytics/working_with_text_data.html).