Week 3: Influencing Thought for Fun or Profit

Nate Bachmeier

TIM-8150: Artificial Intelligence

October 4, 2020

Northcentral University

# Influencing Thought for Fun or Profit

The NCU Political Action Committee (NPAC) seeks to promote its ideology through highly targeted censorship. Modern censorship does not restrict free speech; instead, it increases the noise and drowns the signal (Thomas, 2019). One critical challenge the organization faces is its ability to scale-out personalized communications with potential voters. Traditionally, businesses approach these problems by either hiring armies of people or resorting to mass marketing campaigns. However, NPAC lacks funding to employ a large staff, and modern spam filters reduce email blast effectiveness.

Instead, NPAC chooses to modernize its tactics and focus solely on Natural Language Processing (NLP) and social media graphs. “[…] NLP is an interdisciplinary field [that] studies and develops algorithms and systems, enabling computers to understand and perform tasks involving human language (Sintoris & Vergidis, 2017, p. 135).” NPAC plans to use these technologies to both parse free form text and also produce novel commentary. Maximizing the resonation of custom content with the audience requires a personalized voice. For instance, the manner that people speak in an academic forum differs from Facebook or Twitter. NLP language models can assist in these situations as well by adopting different vocabularies and alternative sentence structures.

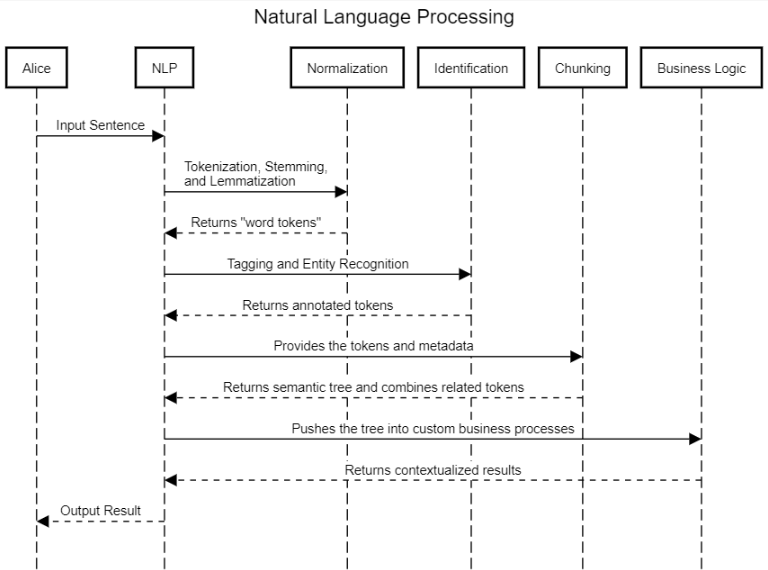
# How does NLP work

Natural Language Processing (NLP) sits at the intersection of artificial intelligence, human language, and computer science.

## Language Parsing

NLP systems typically begin with sentence normalization, combining and annotating tokens, and finally performing custom business logic (see Figure 1) (Edureka, 2018). Using strategies like Lemmatiziation and Stemming enables the parsers to reduce the variability between sentences, such as removing verb-tensing. Next, annotations are associated with the words by subsystems like Named Entity Recognition (NER) that discover the sentence’s critical components. After chunking related tokens together, the scenario-specific business logic can operate on a semantic representation of the text. Depending on the use-case, these steps could be massive subsystems or single lines of code.

Figure 1: NLP Process



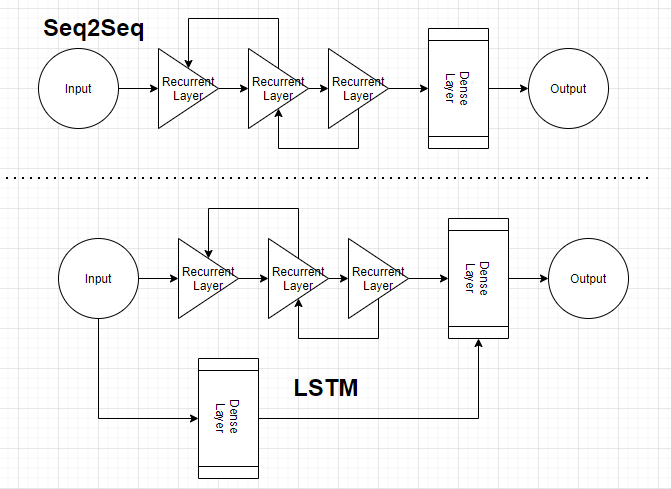
## Deep Learning

NLP appears across a wide range of use cases like language translation, speech-to-text, and sentiment analysis. In biology, animal brains accomplish these tasks through meshes of neurons that transmit signals across connected synaptic (transforming) and activation (filtering) links (Keller, Liu, & Fogel, 2016). Computer scientists mimic this behavior with Deep Learning on Neural Networks, which are essentially weighted graphs. Generally, NLP architectures use Recurrent Neural Network (RNN) structures containing connectivity loops to previous layers (see Figure 2). More advanced designs include subnets for memory retention (see Table 1), encoding and decoding segments, and greater parallelization from attention vectors (Fu, 2019). Researchers and engineers can add or remove these subsystems to optimize for a specific use-case.

Table 1: Example progressions of N.N. architecture complexity

|  |  |
| --- | --- |
| Algorithm | Description |
| seq2seq | Simple Recurrent N.N. (RNN) for a token sequence to sequence prediction. These systems are easy to implement but lack memory |
| Long Short Term Memory | Extends the seq2seq by including a “long term” cache to hold context information |
| Transformers | State of the art solution for massively parallel NLP through attention vectors and position encoding |

Figure 2: Abstract Diagram of Differences



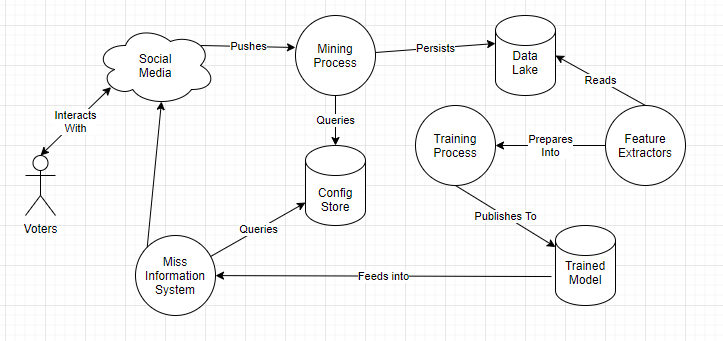
# Resources and Implementation

The publishing pipeline contains distinct features for mining, feature extraction, model training, and custom content generation.

## Ingestion Subsystem

NPAC operates a distributed system that is continuously mining social media platforms and persisting that information into a data lake (see Figure 3). This aspect of the system utilizes WebSockets to stream updates into the big data system. When these results arrive in the data lake, they are semi-structured JSON (JavaScript Object Notation) with numerous opaque strings. Additional details of this subsystem are outside the scope of this paper.

Figure 3: System Design



## Feature Extraction Process

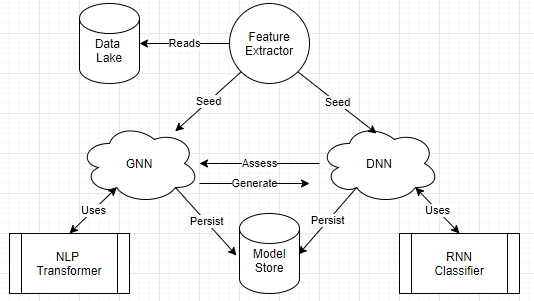
The first steps to any business intelligence problem are identifying the specific questions and locating facts to support answers (Snee, 2015). When researchers ignore this preparation, it produces in garbage-in/garbage-out results. For instance, Alsudias et al. (2014) built an NLP system for predicting where the user was during the submission (e.g., restaurant or night-club). Their approach extracts keywords from Yelp reviews (using term frequency), business metadata (e.g., name and location), and tweet metadata (e.g., timestamp). These features flow into a random forest classifier that determines the user’s location with a 74% accuracy. However, using only the business metadata produces an 88% accuracy, indicating that these additional details provide negative value.

NPAC has specific requirements to model social media users’ speech patterns and then create new content in their voice. The Feature Extraction Process must therefore consider the user’s metadata (e.g., age and locale), the online community properties (e.g., forum name), the posted content, and any quality ratings (e.g., Facebook Likes). There are several considerations to augment this process. For instance, adding a filtration step to remove comments with negative ratings might create more well-liked personalities. However, it could also be advantageous to generate trolls that argue an alternative position, reinforcing NPAC’s position that the other side is illegitimate or less sophisticated.

## Training Process

Around 2014, GAN (Generative Adversarial) Networks became the state-of-the-art approach to produce high-quality fabricated content (Fridman, 2020). These systems utilize a feedback loop between a Generative N.N. (GNN) and Discriminator N.N. (DNN). Each iteration outputs a ‘Deep-Fake’ asset and assesses its likelihood of being legitimate. This process enables both systems to learn from one another, continuously improving. According to Fridman (2020), it an arms race to detect Deep Fakes because any advances in DNN naturally improve GNN results. NPAC leverages this methodology for self-teaching its systems to deliver more accurate content (see Figure 4). The organization’s solution uses the NLP transformer to improve parallelization over LSTM and a second RNN classification network. During the training process, periodic snapshots archive the content and model state for offline troubleshooting use cases.

Figure 4: Training Configuration



## Inference Process

The last stage of NPAC’s misinformation system is a crawler that pushes its message across social media. Before publishing content, the system first collects metadata about the specific discussion thread, the participant’s demographic details (e.g., age group), and summarizing the most liked comments. These features seed the generation process to improve its relevance with the specific conversation. When choosing the content’s voice, the inference process uses basic statistics to determine the median group.

Today, summarizing comments uses a term frequency distribution to seed the GNN model and improve relevance. Boorugu and Ramesh (2020) describe state-of-the-strategies that produce abstracts instead of extracts. While NPAC is not actively investing in this area, it does open doors for future innovation. For instance, by sharing articles with the automated abstract, can the fake accounts gain more followers?

# Future Directions

Two areas the organization wants to pursue next include empathy and interactive communication. While there are likely to be many additional improvements, such as stylistic systems, these two areas are most relevant to NPAC’s business model.

## Emotional Awareness

Huang et al. (2019) state that incorporating human emotion into A.I. systems is a decade away. Perhaps this is true for the general case, but initial wins also exist along the way. For instance, the system could also capture sentiment analysis information about the thread before commenting. Those reactions can then feed into a stylistic decision model that chooses more appropriate tones. This capability would allow the publishing pipeline to emit more impactful content that harnesses the voter’s mental state.

## Interactive Communications

Another critical limitation of mass propaganda comes from being a static monologue versus a dialog. For instance, a post today might amount to “vote for XYZ,” versus tomorrow's post allows viewers to ask “why” or engage in debate. In 1950, Turing proposed chatbots, but only a handful have thrived since (Singh & Thakur, 2020). NPAC needs to identify iterative steps to increase interactive capabilities to continue scaling out the message without requiring additional humans. Perhaps scoping the data domain and following a rule engine begins the journey, but it needs to become more than that.

# Conclusions

NPAC uses natural language processing to remain a lean organization without impacting its ability to deliver personalized propaganda across social media. Open-source software like Keras and Tensorflow make these advanced capabilities available even to small organizations. This ease of access also helps researchers devise sophisticated deep learning architectures that retain more context and learn more complex rules.

Next, these concepts coalesce in a reference system design that mines social media graphs. Those data points connect into a GAN to produce deep fake content that is self-assessing and becoming more accurate with each iteration. NPAC then uses those models to match online conversations’ language and tone to inject its messages covertly. Despite the system being in infancy, there is much potential for this program anywhere that voter manipulation is necessary.

# References

Alsudais, A., Leroy, G., & Corso, A. (2014). We Know Where You Are Tweeting. *IEEE International Congress on Big Data* (pp. 594-600). Anchorage, AK: The Institute of Electrical and Electronics Engineers. doi:https://doi-org.proxy1.ncu.edu/10.1109/BigData.Congress.2014.91

Boorugu, R., & Ramesh, G. (2020). A survey on NLP based text summarization for summarizing product reviews. *International Conference on Inventive Research in Computing Applications*, (pp. 352-356). Coimbatore, India. doi:https://doi-org.proxy1.ncu.edu/10.1109/ICIRCA48905.2020

Edureka. (2018, October 16). *Natural language processing in 10 minutes*. Retrieved from YouTube: https://www.youtube.com/watch?v=5ctbvkAMQO4

Fridman, L. (2020, January). *Deep Learning State of the Art*. (Massachusetts Institute of Technology (MIT)) Retrieved from YouTube: https://youtu.be/0VH1Lim8gL8

Fu, Z. (2019). An introduction of deep learning based word representation applied to natural language processing. *International Conference on Machine Learning, Big Data and Business Intelligence*, (pp. 92-104). doi:https://doi-org.proxy1.ncu.edu/10.1109/MLBDBI48998.2019.00025

Huang, M., Rust, R., & Maksimovic, V. (2019). The feeling economy: managing in the next generation of artificial intelligence. *California Management Review, 61*(4), 43-65. doi:https://doi-org.proxy1.ncu.edu/10.1177/0008125619863436

Keller, J., Liu, D., & Fogel, D. (2016). *Fundamentals of Computational Intelligence.* John Wiley & Sons.

Singh, S., & Thakur, H. (2020). Survey of various A.I. chatbots based on technology used. *International Conference on Reliability, Infocom Technologies and Optimization*, (pp. 1074-1079). Noida, India. doi:https://doi-org.proxy1.ncu.edu/10.1109/ICRITO48877.2020.9197943

Sintoris, K., & Vergidis, K. (2017). Extracting busines process models using NLP techniques. *IEEE 19th Conference on Business Informatics* (pp. 135-139). Thessaloniki, Greece: Institute of Electrical and Electronics Engineers. doi:https://doi-org.proxy1.ncu.edu/10.1080/014492997119879

Snee, R. (2015). Practical approach to data mining. *Quality Engineering, 27*, 477-487. doi:10.1080/08982112.2015.1065322

Thomas, R. (2019). The new era of NLP. *Scientific Computing with Python.* Austin, Texas: SciPy. Retrieved from https://youtu.be/KChtdexd5Jo