Project Description

Elevator Pitch

The Customer: Neolth de-escalates stress and mental health crises for high school and college students across the US. Neolth is provided to students as a school health service, after being purchased by the school. There has been a 30% increase in demand for campus counseling center appointments since 2010 (Winerman, 2017; Reetz et al., 2016). Counseling centers have struggled to meet this demand, with over one-third of schools reporting a waitlist with an average of 17.7 business days (LeVinees et al., 2018). Even without waitlists, it is impossible for counselors to manually monitor student mental health on a continuous basis. That's why over 40% of schools do not offer triage appointments (LeVinees et al., 2018). Unfortunately, that means many students cannot get care when they need it. With COVID-19 campus closures and remote learning throughout the country, now more than ever counseling centers require an automated, continuous monitoring system that will allow them to track student mental health remotely and triage students into care as needed.

The Value Proposition: The adoption of the Neolth digital health platform will provide counseling centers continuous mental health monitoring, benefiting schools by preventing inadequacies in care and potential lawsuits (KCRG, 2019). Adoption of Neolth's predictive model coupled with digital phenotyping techniques will make the process of identifying students at risk for suicide faster, easier, and more accurate for schools. Adoption of Neolth's personalized behavioral plans to manage stress and promote mental wellbeing as an adjunct to counseling center care will increase immediate access to care for students. Mass adoption will benefit schools, students and their families by providing a cost-effective way to increase the overall standard of student mental healthcare.

The Innovation: The predictive model will identify the student-users who are suicidal and require immediate professional care. As preliminary work, a state-of-the-art model in Natural Language Processing (NLP) was built to identify suicidal student-users through written text data. The generalized performance of this prediction model was validated using social media posts of adolescents and young adults. To transform this model from a functional prototype to a commercially viable and life-saving product, additional research and development is needed. In this Phase 1 SBIR, we will combine artificial intelligence and human intelligence by incorporating active learning into the model to improve its performance. We will also include digital phenotyping components to identify students who do not disclose their suicidal ideation through text data on the Neolth application, but rather exhibit digital behaviors consistent with suicide risk.

Innovation

Neolth's automated continuous monitoring system will replace the manual monitoring performed by school and university counseling centers by automatically flagging students at risk of committing self-harm or suicide. This multi-label classification predictive model will meet the most urgent needs of school clinicians and counselors. The functioning of the predictive model has been successfully prototyped by implementing the latest technology: a pre-trained unsupervised Natural Language Processing model named BERT –

Bidirectional Encoder Representations from Transformers (Horev, 2018; Sileo, 2019; Vig, 2018). This model will be applied to students' written text data in Neolth's application for suicide risk analysis. The results generated by the model will be used to flag student-users and alert school mental health counselors that clinical care is needed. The original model has been proven to achieve state-of-the-art performance on several Natural Language Understanding tasks. Although the evaluation phase of developing the predictive model has been successfully completed with a reasonable model accuracy percentage (84%), this prototype requires refinement to improve accuracy and be commercially viable.

Phase I of the project will focus on improving the accuracy of the functional prototype, which will ultimately help correctly classify students at risk for suicide. The system must accurately measure suicide risk, in order to send appropriate alerts to the counseling center so they can make sound treatment decisions. However, there are possibilities that the model will not correctly classify the text data under the right labels. This is when human intelligence plays an essential role. Through the concept of Active Learning (Paolo Tamagnini, 2019; Rosaria Silipo, 2019), which is an iterative process of gathering expert feedback from clinicians about the model's performance, a high-performance classifier will be developed. Additionally, Digital Phenotyping will be implemented to investigate behavioral patterns indicative of suicidal ideation beyond explicit keywords (Rooksby et al., 2019; Liang et al., 2019). The merger of Cognitive Analytics (Linguistic Analytics) and Behavioral Anomaly Detection (Pattern-based Analytics) is a strong framework to accurately assess mental health. This complementary Multimodal Learning approach will de-risk the ability to accurately flag students with suicidal ideation. Significant research and development in software is needed to advance Neolth's triaging system from a functional prototype to a life-saving process.

In Phase II, the model will be further developed by improving the user interface and integration of this alert system into the school counseling center's workflow. Real-time testing with counseling centers will be conducted to optimize the performance of the alert system.

Company and Team

Neolth de-escalates student stress and health crises by empowering students to develop the resiliency skills needed to take control of their health and future. Our digital platform offers personalized emotional support and continuous mental health monitoring, reducing the burden of high care demands on high school and university counseling centers. Neolth's end-to-end solution is already on the market and includes assessment, personalized treatment recommendation, treatment completion, and progress tracking. Neolth is used as an adjunct to in-person care, as part of the Stepped Care Model (Centre in Innovation in Campus Mental Health, 2020).

Neolth is led by an experienced team and further supported by a clinical and business advisory board affiliated with Oxford University, Stanford University, Harvard Business School, University of California Berkeley, University of California San Francisco, Pear Therapeutics, Carnegie Mellon University, the University of Pittsburgh Medical Center, the Massachusetts Institute of Technology, and more (Neolth, 2020). CEO and Co-founder Katherine Grill, PhD is a Neuroscientist specializing in Integrative and Behavioral Health. She has an MA in Psychology and BS in Art Therapy. She will be the PI and serve as a clinical expert on this project. Before founding this company in 2018, she taught psychology and precision neuroscience as a university professor for 7 years. She is trained extensively in mind-body and creative arts

therapies, has conducted behavioral care in inpatient and outpatient psychiatric units, a substance abuse rehabilitation center, and in a hospital-based palliative care setting. She is also a certified Advance Care Planning Facilitator. While working at Children's National, one of the top ten pediatric hospitals in the US, Katherine specialized in adolescent and young adult behavioral health. She worked on three NIH-funded R01 clinical trials. She also co-founded a stress management group for young adults with chronic illness and was part of the effort to build an Integrative Medicine Department at Children's National.

Claire Wheeler, Neolth's CMO, has an MD in Emergency and Integrative Medicine and PhD in Clinical Psychology. She will serve as a clinical expert on this project. Claire began her career as a licensed clinical psychologist and physician specializing in stress and trauma treatment. After practicing Emergency Medicine, she completed a fellowship in Integrative Medicine at the University of Arizona. She was the Principal Investigator of a 4-year study of the emotional, social, and psychological effects of traumatic injury and stress at the University of Michigan and the Legacy Health System in Portland. She also completed research in the 80s on nuclear magnetic resonance (NMR) spectroscopy for measuring metabolic activity, that was foundational to the development of MRI. She went on to found MindBody Medicine of Portland, a private practice where she provided integrative medical services to her patients. Claire has been an Assistant Professor at the OHSU/PSU School of Public Health in Portland, Oregon for over 15 years, where she teaches graduate and undergraduate classes in health behaviors, integrative medicine, and stress. She is the author of *10 Simple Solutions to Stress* and *Pocket Guide to Stress Management*. She was also key personnel for a grant from NCCAM (now NCCIH) to Oregon Health & Science University to integrate mind-body medicine and other integrative practices into the curriculum of the medical and nursing schools. Claire has been a Clinical Supervisor and Senior Faculty Member at the Center for Mind Body Medicine, a role she's held for almost 20 years. In that role, Dr. Wheeler trained hundreds of health professionals from all over the world in mind-body medicine and how it can be integrated into clinical medical and mental health practices as well as community health initiatives. Combined, Katherine and Claire have over 40 years of clinical experience in adolescent and young adult behavioral health.

Chia-ying Lee has a PhD from CSAIL MIT in machine learning and artificial intelligence. She will serve as the Program Director. She has expertise in natural language processing and was the former Lead Data Scientist at Virta Health, in charge of building machine learning algorithms for their personalized diabetes management solution. She has further experience building inference algorithms, deep learning and recurrent neural networks, speech recognition, computational cognitive science, and HMM-based synthesizers. She has conducted research at Google, Microsoft, MIT, and USC Information Sciences Institute. Dr. Lee has served on the advisory board of Neolth for the past two years. Dr. Lee will provide support and oversight to the Neolth team in building and refining the advanced analytics for suicide risk detection.

Woojin Park is pursuing an MS in Business Intelligence & Data Analytics at Carnegie Mellon University specializing in Data Science. He has a BS in Information Systems and MS in Data Science from Syracuse University. He has expertise in Big Data analytics, Natural Language Processing and Computer Vision. He was a former Big Data Analyst at Kia Motors, led various data analytics projects including Inventory Control, Time-Series Modeling and Cost Optimization over 4 years. He was formerly a data science researcher in Syracuse University, building models to assess the importance of different location nodes applying a PageRank algorithm. He is currently a Data Scientist and Project Manager at Students Using Data For Social Good (SUDS) at Carnegie Mellon University, leading a machine learning project in

partnership with the Non Profit Organization - Family Independence Initiative (FII), CA. Woojin is also a Data Scientist Intern at Neolth.

Nidhi Bhaskar is pursuing her Masters in Information Systems Management and specializes in Data Science and Analytics at Carnegie Mellon University (CMU). She formerly worked as an Application Developer Associate on several Business Intelligence Analytical projects at Accenture. Through academic explorations at CMU, she successfully executed projects in the domain of Statistical Analysis, Machine Learning, and Natural Language Processing. She currently serves as the Data Project Chair of Students Using Data for Social Good (SUDS) at CMU and is adept at working in an agile environment, evaluating customer requirements, and building a strategic perspective. Nidhi is also a Data Scientist Intern at Neolth.

Woojin and Nidhi will work as Data Scientists on this project.

Lukas Grill, Neolth's COO and Co-founder, has a BA in Business & Economics and an MS in Engineering. His specialities are business development and IT project management, with 10+ years of experience in both fields. He joined the Project Management Office of an international supplier of intelligent road traffic telematic solutions at the company's headquarters in Vienna in 2012. He transferred to the North American office in 2014 to deliver large scale IT projects on the west coast. His portfolio includes managing the software development for the Golden Gate Bridge tolling system. As the Associate Vice President of a US IT company, he managed the finances and timeline of large-scale ($15M) technical projects as well as 30 employees. Lukas will serve as the Project Manager on this grant, ensuring timeline and budget oversight.

Technical Discussion and R&D Plan

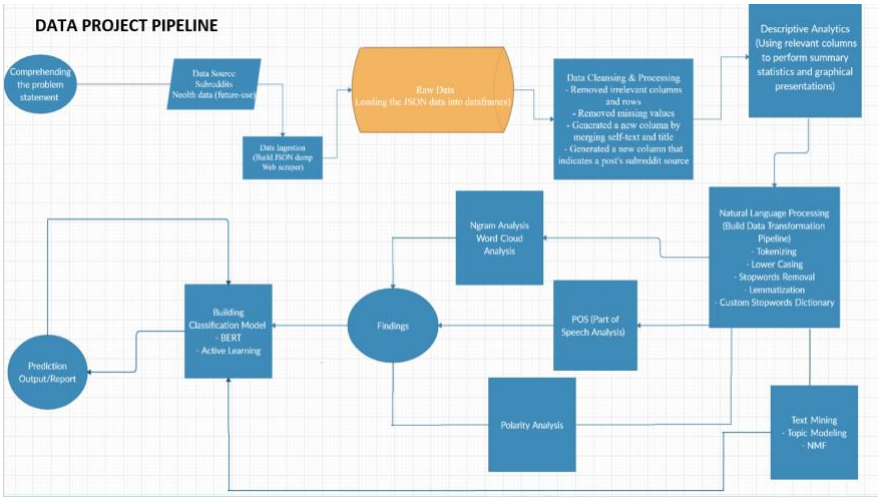
Background - BERT Model: In order to create a highly-accurate multi-label classification model, we will need to conduct significant research and development. The underlying technology on which we wish to implement the predictive model is BERT, a state-of-the-art language model developed by researchers at Google AI language. BERT, which stands for Bi-directional Encoder Representations from Transformers, is a Natural Language Processing (NLP) technique created and published in 2018 (Horev, 2018). The large transformer, which is a neural-network based model for NLP pre-training, is capable of performing context-based analysis. This overcomes the disadvantages of other NLP context-free techniques such as word2vec or Global Vectors for Word Representation (GloVe). This transformer adopts an attention mechanism that takes a token from each input sequence and focuses on any other token (Horev, 2018). This facilitates the learning of contextual relations between the words in the text data. The advantage of using this pre-trained model is the bi-directional nature of the transformer. Unlike the directional models which take in text data sequentially and process the data either from left-to-right or right-to-left, the bidirectional transformer processes all the data at once. This characteristic allows the model to analyze data in a contextual manner by learning its meaning based on the surrounding words. This characteristic is key to the success of our suicidal risk analysis, as it is not just keywords but overall contextual meaning that must be analyzed.

BERT allows us to extract high quality language representations from a large text data set and then fine-tune the representations for our specific task. Technically, we will ensemble two main learning frameworks: pre-training (Unsupervised task) and Fine-Tuning (Supervised task) with labeled data specific to the task. The BERT model is pre-trained with two corpuses: BooksCorpus (800M words) and English Wikipedia (2,500M words). The pre-training step allows the model to learn general word representations of the English language and also reduces the need for many heavily-engineered task-specific architectures. BERT uses multi-layers of attention, incorporates multi-attention heads in each layer, and it supports over a 100

million parameters. These factors contribute to a successful, high-efficiency model which we plan on constructing using BERT. Various hyper-parameters are optimized, which are the set of variables defined before training and before the network structure is designed. To find out the most optimal configuration of parameters such as learning rate, batch size, number of epochs and optimizer, a grid search hyper-parameter tuning function will be employed.

Considering all these attributes, we are confident that a predictive model based on BERT will yield optimized results. By accumulating real-time textual data from Neolth's application and integrating model performance feedback from clinical experts, we can continuously improve the quality and performance of the model. Beyond flagging users vulnerable to committing suicide or self-harm through text analysis, other pattern detection models indicative of suicidal ideation, such as the Digital Phenotyping of users' sleep patterns, app usage or texting style would be deployed. The combination of Cognitive Analytics (Linguistic Analytics) and Behavioral Anomaly Detection (Pattern-based Analytics) forms a strong framework to accurately assess mental health.

Preliminary Work - Proof of Feasibility:



Milestone 1: Collect and Create a Data Set for the Model.

Task 1: Data Scraping. To construct a predictive model, it is necessary to ensure that the input data fits the model perfectly. We built a recursive and automatic Web-scraper for collecting unstructured social media posts from the online platform Reddit, to serve as our base data set (Reddit, 2020). Reddit consists of various 'Subreddits', which are online forums that allow people to post their feelings or experiences on specific topics. Our scraper gathered data from four different subreddits called 'SuicideWatch', 'Depression' 'Happy' and 'Self-Improvement'. These subreddits consist of more than 100,000 posts created by users over the course of a few years, making it easy for us to generate a large dataset containing unstructured text data. We applied Pushshift API to gain dynamic access to collect large dataset recursively as a JSON dump format, aggregate the scraped data into columnar-base dataframe and read it into our system using Pandas Library from Python.

Task 2: Data Preprocessing and Cleaning. After we successfully scraped and transformed the data into the required Python dataframe format, we performed the necessary data preprocessing and cleaning steps. This included the following essential techniques that we aim at implementing to ensure NLP model accuracy: removal of irrelevant columns, treating missing data and outliers, and adding specific columns that are required for our analysis. By incorporating various Feature Engineering techniques and ensuring

that the text data is formatted in the best possible manner, we ensured that the data is of high quality and is ready to be used for various NLP tasks. The clean data then needed to undergo the transformation phase which included the following: tokenizing, lower casing, lemmatizing and stop-words removal.

Milestone 2: Exploratory Data Analysis

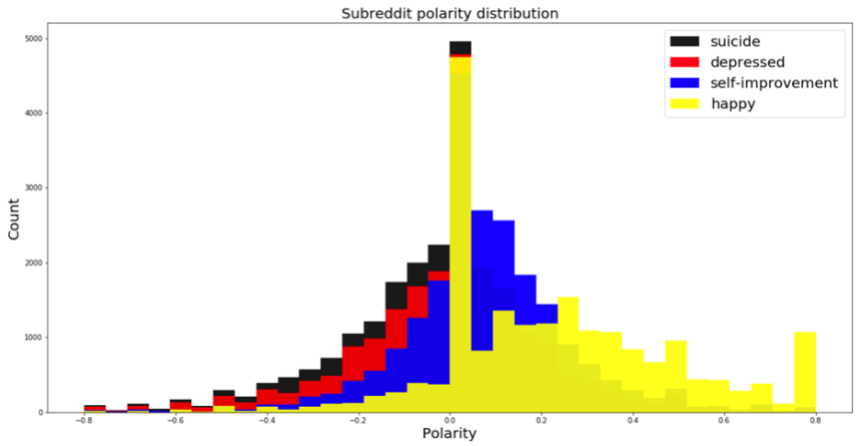
Task 1: **POS – Parts of Speech Analysis**. POS Analysis generates grammatical tagging of each word in a text corpus as corresponding to a particular part of speech based on both its definition and context.

Task 2: **N-gram Analysis** is performed to find a contiguous sequence of N-items from a given Subreddit post. The items of N-gram analysis can be phonemes, syllables, letters, words or base pairs according to the application. The analysis is mainly used to understand which nouns, adjectives, or phrases are most frequently used in each Subreddit group. This analysis helped us materialize and capture each Subreddit group's main keywords and topics. This also helped us come up with accurate labels to classify our text data under the Multi-label Classification model.

Task 3: **Word Cloud Analysis**. Once we found the most frequently appeared words for each subreddit, we applied the Word-cloud library from Python to visualize our findings more efficiently.

Task 4: **Topic Modeling using NMF**. The Multi-Label Classification Model requires that we identify the correct labels under which we would classify the data. Using NLP techniques such as Topic Modeling helped us identify the distribution of the text data under the defined number of labels. We implemented the Non-Negative Matrix Factorization (NMF) Topic modeling to our task which is an 'unsupervised' feature extraction technique. It helped us find out the best latent topics from the large data set. Technically, NMF decomposes high-dimensional vectors into a lower-dimensional representation iteratively until the approximation error converges or the specified number of iteration is reached. The output of NMF modeling was 'Document-Topic' Matrix and 'Topic-Term' Matrix. The use of these techniques coupled with clinical expert opinions and our BERT based classification model's output will allow us to come up with the right labels to correctly classify each post.

Task 5: **Polarity Analysis** is one of the well-known Sentiment Analysis techniques, which is performed to determine the attitude or emotion of the writer i.e., whether it is positive, negative or neutral. The sentiment function of Python library Textblob returns 'polarity' which is the number lines in the range of [-1,1] where 1 means a positive statement, -1 means a negative statement and 0 means a neutral statement. According to our findings, 'SuicideWatch' Subreddit posts show the most negative sentiment distribution and the 'Happy' Subreddit posts show the most positive sentiment distribution.



Milestone 3: Training, Validation and Testing

Task 1: **Training.** First, our BERT based predictive model was fit on a training dataset. Once the input vector was trained, the result was compared with the target value. Based on the result of this output comparison, the parameters of the model are adjusted. This is called model fitting.
Task 2: Validation. The fitted model was then used to predict the result of the input vector in the validation dataset. This dataset provided an evaluation of our model fit on the training data. We implemented the 'Early Stopping' functionality, which stops the training when the error/loss on the validation dataset increases by more than 2 or 3 epochs consecutively. This technique helped us avoid overfitting issues.

Task 3: Testing. Finally, we performed a final model evaluation using the unseen test dataset. The evaluation matrix for our Multi-label Classification model consisted of (1) 83.9% Accuracy (Number of the input cases that are correctly identified as either true positive or true negative out of the total number of cases) and (2) F1-Score of 83.8% (weighted average of the precision and the recall, which is the score that takes both false positive and false negatives into account. It is usually more useful than accuracy when we have an uneven class distribution).

Results of preliminary work:

In [54]:
```
precision, recall, fscore, support = score(y_test, predicts)
target_names = ['SuicideWatch', 'depressed', 'happy', 'selfimprovement']
print(classification_report(y_test, predicts, target_names=target_names))


f1_score(y_test, predicts, average='weighted')
```

```
                precision    recall  f1-score   support

  SuicideWatch       0.80      0.80      0.80      1562
     depressed       0.73      0.72      0.72      1365
         happy       0.93      0.94      0.93      1534
selfimprovement       0.89      0.89      0.89      1539

      accuracy                           0.84      6000
     macro avg       0.84      0.84      0.84      6000
  weighted avg       0.84      0.84      0.84      6000
```

Out[54]: 0.8388372854791574

All preliminary work for the proof of concept was completed by Woojin and Nidhi and supervised by Drs. Lee and Grill. Although the base model is completed, significant research and development is needed to progress Neolth's suicidal risk analysis system from a functional prototype to a life-saving process. In Phase 1 of this SBIR grant, the following milestones would be achieved to transform the functional prototype to a viable commercial product:

Milestone 1: **Improving model's accuracy through hyper-parameter tuning and active learning.**
Although our base model has an accuracy rate of 84%, we need to improve this so the model can be commercially viable. Active learning is a special case of Machine Learning in which a learning algorithm can interactively query a subject matter expert (clinician or school counselor) to label new data points with the desired outputs.

Task 1: Clinical experts review/label posts. Text data from posts on Neolth will be sent for review to the clinicians who will classify the posts to the correct clinical label: (1) Suicidal, needs immediate action; (2) Depressed, needs clinical care within the week; (3) No clinical action needed.

Task 2: Match the model's output with correct labels. Once we have collected the labeled posts from the clinicians, we would run the same data on the predictive model. The model's output will be compared to the results gathered from the clinicians. This process will help the model determine the difference between students who are suicidal and need immediate care versus students who are depressed but are not

a suicide risk. This is an essential component that must be tested and developed for the model to function effectively without false negatives or false positives.

Task 3: Incorporate clinical feedback into the model. Once the misclassified data is placed under the right labels with the help of clinicians, we would re-train the model. This would improve the accuracy of the model and its efficiency in dealing with similar ambiguous text data thereafter.

Team and Timeline: Milestone 1 will be overseen by Dr. Lee. Nidhi and Woojin will identify the instances of misclassification needing review. Drs. Grill and Wheeler will review classification errors and provide clinical feedback, which Nidhi and Woojin will incorporate into the model. We expect completion to take 8 weeks.

Milestone 2: Research and develop digital phenotyping components. To overcome the challenges of identifying the 20% of students who refrain from disclosing their suicidal thoughts, we will implement Digital Phenotyping (Rooksby et al., 2019; Liang et al., 2019). Digital Phenotyping is a fairly new concept, so considerable research and development is required to create an accurate and commercially viable suicide risk detection system.

Task 1: Identifying feasible behavioral patterns to track. Significant research is necessary to identify the type of digital behavioral patterns best suited for representing suicide risk in students. We believe that, with the help of Anomaly Detection (machine learning technique to identify data that does not conform to the standard patterns), we will be able to track and flag students using Digital Phenotyping.

Task 2: Identifying acceptable sources of data. Various methods have been suggested, including camera, microphone, location, and keyboard usage. Important to note is a recent study from the UK (Rooksby et al., 2019), which reviewed students' acceptability of various types of digital phenotyping. Students felt some of these approaches (camera, microphone) were an invasion of privacy and preferred using data from the accelerometer, light sensor, tracking of non-dating apps, and tracking the number of messages sent and received (not content). As part of the Phase 1 proposal, we would conduct focus groups with the 43,000 students and counting affiliated with Neolth schools, to understand their attitudes towards Digital Phenotyping data sources.

Task 3: Use focus group results to inform model design. The focus group results will provide an understanding of the extent to which students are willing to share their data. The team will then decide the best way to add acceptable data sources into the model, considering clinical and technical perspectives. This complementary multimodal analysis will help identify students who are unwilling to express their suicide ideation through text data, thereby improving the commercial viability and life-saving potential of our solution.

Technical Challenges: Although creating the proposed alert system to flag students suffering from suicidal ideation and depression is possible and promising, the most demanding technical challenge is avoiding misclassifying students who need immediate care. This would be due to ambiguous text data (lack of indication of which label it belongs to) which the model would fail to recognize or due to students refraining from expressing suicidal ideation through text data. While digital phenotyping could mitigate this risk, as a new field it has its own challenges: adoption of students due to privacy concerns and incorporating data into the model while maintaining the accuracy of the algorithms that have been previously demonstrated (Rooksby, 2019). Both false negatives and false positives pose risks to implementation that need to be addressed through further research and development to build a highly accurate model that is commercially viable.

A significant technical challenge is to avoid misclassification that leads to false negatives or false positives. While implementing Active Learning as part of Milestone 1 would help with this risk, there is the challenge of inter-rater variability. It is possible that clinicians at our school customers will classify students' posts differently than Drs. Grill and Wheeler. We would need to determine how to integrate feedback into the model when it is conflicting, and how each customer decides if a student needs immediate care. Furthermore, from the Polarity Analysis graph displayed above, it is evident there is an overlap in the distributions between depression and suicide. It can be difficult to distinguish between students who are depressed and withholding thoughts of suicidal ideation (which happens frequently in this age range) and students who are depressed but not suicidal (Rooksby et al., 2019). Such misclassification would give rise to an increase in the number of false positives and false negatives. False positives would add burden to the school counseling center and could upset the student, while false negatives could lead to the schools missing out on students who are suicidal.

Another significant technical challenge is identifying the 20% of students who do not give clear warning signs or disclose their suicidal ideation (Jason Foundation, 2020). The system must accurately measure the suicide risk in order to send alerts to the counseling center so that they can make sound treatment decisions. However, since Digital Phenotyping is a new field, there is limited information about the validity of data sources that can be used to identify depression and suicidality. Furthermore, students have privacy concerns about sharing their data with a technology company and with their school (Rooksby et al., 2019).

Lastly, since BERT is a relatively new model, we need to understand how the selected combination of hyper-parameters, the optimizer and the designed architecture influence our outcomes. The design of the neural network architecture decides the success of the model. As the hyper-parameters and architecture of the model changes, the accuracy varies. This variance leads to a potential challenge we expect to face, which is the identification of the most suited combination of parameters that would work the best for our predictive model. In most situations, increasing the training data set adds diversity and it decreases the generalized error because the model becomes more general by virtue of being trained on more examples. However, because of the limited computational power of the local machine, we need to build and train our model on a cloud platform using GPU/TPU at AWS and Google Colab. This would result in various potential problems such as Python Library Versioning, Code compatibility over platforms and Data Quality management. To mitigate these various uncertainties in transitioning to the cloud server, our team would need to build new IT governance and control processes.

All of the elements discussed above make up the components of the suicide risk detection system. Each element must be functioning high enough so that this system can be integrated into clinical care at high schools and universities. This poses a significant technical challenge to reduce the amount of false negatives and false positives to provide accurate clinical alerts. It is essential that these technical hurdles be cleared so that the 7.3 million US students experiencing suicidal ideation can benefit from the advantages of this technology (Liu, 2018).

References

[1] Winerman, L. (2017). By the numbers: Stress on campus, college and university counseling centers are seeing an uptick in the number of students seeking help. *American Psychological Association, 48*(8), pg. 88.

[2] Reetz, D.R., Bershad, C., LeViness, P., & Whitlock, M. (2016). The Association for University and College Counseling Center Directors (AUCCCD) Annual Survey. Retrieved from: https://www.aucccd.org/assets/documents/aucccd%202016%20monograph%20-%20public.pdf

[3] LeVinees, P., Bershad, C., Gorman, K., Braun, L., & Murray, T. (2018). The Association for University and College Counseling Center Directors (AUCCCD) Annual Survey - Public Version. Retrieved from https://www.aucccd.org/assets/documents/Survey/2018%20AUCCCD%20Survey-Public-June%2012-FINAL.pdf

[4] KCRG.com (2019). Family suing Iowa for their son's suicide. Retrieved from: https://www.kcrg.com/content/news/Family-suing-Iowa-for-their-sons-suicide-512265282.html

[5] IBIS World(a). (2020). Colleges and Universities in the US: Market size 2004 - 2026. Retrieved from: https://www.ibisworld.com/industry-statistics/market-size/colleges-universities-united-states/#:~:text=The%20market%20size%2C%20measured%20by,is%20%24528.4bn%20in%202020.&text=The%20market%20size%20of%20the%20Colleges%20%26%20Universities%20industry%20in%20the,average%20between%202015%20and%202020.

[6] IBIS World (b). (2020). Private schools in the US industry statistics. Retrieved from: https://www.ibisworld.com/united-states/market-research-reports/private-schools-industry/

[7] IBIS World (c). (2020). Public schools industry in the US: Market research report. Retrieved from: https://www.ibisworld.com/united-states/market-research-reports/public-schools-industry/

[8] US Department of Education. (2020). Allocations for Section 18004(a)(1) of the CARES Act. Retrieved from: https://www2.ed.gov/about/offices/list/ope/allocationsforsection18004a1ofcareSact.pdf

[9] DeAngelis, T. (2019). College students' mental health is a higher priority: Colleges and universities are increasing efforts to address student mental health. Retrieved from: https://www.apa.org/monitor/2019/12/numbers-college

[10] Forestieri, K. (2019). State launches $50M program for school-based mental health: Bevy of recent bills boost the role of school districts in health care, suicide prevention. Retrieved from: https://www.paloaltoonline.com/news/2019/08/02/state-launches-50m-program-for-school-based-mental-health#:~:text=The%20California%20State%20Legislature%20is,illness%20through%20campus%2Dbased%20Services.

[11] U.S. Department of Education. (2019). U.S. Department of Education Announces New Grant Awards to Address School Safety and Improve Access to Mental Health Services. Retrieved from: https://www.ed.gov/news/press-releases/us-department-education-announces-new-grant-awards-address-school-safety-and-improve-access-mental-health-services

[12] Gordon, L. (2019). California colleges expand mental health services to meet rising needs. Retrieved from: https://edsource.org/2019/california-colleges-expand-mental-health-services-to-meet-rising-needs/610151

[13] American College Health Association (ACHA) National College Health Assessment (NCHA). (2018). Undergraduate student reference group. Retrieved from: https://www.acha.org/documents/ncha/NCHA-II_Fall_2018_Undergraduate_Reference_Group_Executive_Summary.pdf

[14] Burrel, J. (2020). College and teen suicide statistics. Retrieved from: https://www.verywellmind.com/college-and-teen-suicide-statistics-3570768

[15] NICHE. (2019). The top 50 boarding schools in New England for 2019. Retrieved from: https://www.niche.com/blog/top-50-boarding-schools-in-new-england/

[16] Blackstone Launchpad by Techstars. (2020). Retrieved from: https://www.blackstonelaunchpad.org/

[17] UCSF Innovation Ventures. (2020). Startup 101. Retrieved from: https://innovation.ucsf.edu/startup-101

[18] Centre for Innovation in Campus Mental Health (2020). Stepped Care Approach. Retrieved from: https://campusmentalhealth.ca/toolkits/campus-community-connection/models-frameworks/stepped-care-model/

[19] Horev, R. (2018). Bert Explained: State of the art language model for NLP. Retrieved from: https://towardsdatascience.com/bert-explained-state-of-the-art-language-model-for-nlp-f8b21a9b6270

[20] Sileo, D. (2019). Understanding BERT Transformer: Attention isn't all you need. Retrieved from: https://medium.com/synapse-dev/understanding-bert-transformer-attention-isnt-all-you-need-5839ebd396db

[21] Vig, J. (2018). Deconstruction BERT: Distilling 6 Patterns from 100 Million Parameters. Retrieved from: https://towardsdatascience.com/deconstructing-bert-distilling-6-patterns-from-100-million-parameters-b49113672f77

[22] Bishan Yang, Jian-Tao Sun, Tenjiao Wang, Zheng Chen (2009). Effective Multi-label Active Learning for Text Classification. Retrieved from: https://www.microsoft.com/en-us/research/wp-content/uploads/2009/01/sigkdd09-yang.pdf

[23] Rooksby, J., Morrison, A., & Murray-Rust, D. (2019). Student perspectives on digital phenotyping: The acceptability of using smartphone data to assess mental health. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-14).

[24] Liang, Y., Zheng, X., & Zeng, D.D. (2019). A survey on big-data driven digital phenotyping of mental health. Retrieved from: https://www.sciencedirect.com/science/article/pii/S1566253518305244

[25] Neolth Inc. (2020). About Us. Retrieved from: https://www.neolth.com/about-us

[26] Alammar, J. (2018). The Illustrated BERT, ELMO, and co. (How NLP cracked Transfer Learning). Retrieved from: http://jalammar.github.io/illustrated-bert/

[27] Rosebrock, A. (2019). Rectified Adam (RAdam) optimizer with Keras. Retrieved from: https://www.pyimagesearch.com/2019/09/30/rectified-adam-radam-optimizer-with-keras/

[28] Reddit. (2020). Retrieved from: Reddit.com

[29] Susan, L. (2019). Anomaly Detection for Dummies. Retrieved from: https://towardsdatascience.com/anomaly-detection-for-dummies-15f148e559c1

[30] The Jason Foundation. (2020). Youth suicide statistics. Retrieved from: http://prp.jasonfoundation.com/facts/youth-suicide-statistics/

[31] Liu, C.H., Stevens, C., Wong, S.H.M., Yasui, M., & Chen, J.A. (2018). The prevalence and predictors of mental health diagnoses and suicide among U.S. college students: Implications for addressing disparities in service use.

Depression and Anxiety, *https://doi.org/10.1002/da.22830*.

[32] Jonathan, Hui. (2019). Word Embedding & Glove
Retrieved from:
https://medium.com/@jonathan_hui/nlp-word-embedding-glove-5e7f523999f6

[32] Paolo Tamagnini and Rosaria Silipo(2019). Labeling with Active Learning
Retrieved from:
https://www.knime.com/blog/labeling-with-active-learning

Potential References: https://time.com/5190291/anxiety-depression-college-university-students/#:~:text=The%20average%20university%20has%20one,International%20Association%20of%20Counseling%20Services.

https://www.sprc.org/colleges-universities/consequences#:~:text=Mental%20health%20problems%20can%20affect%20a%20student's%20energy%20level%2C%20concentration,%2C%20and%20optimism%2C%20hindering%20performance.&text=Research%20suggests%20that%20depression%20is,anxiety%20can%20increase%20this%20association.

https://healthymindsnetwork.org/wp-content/uploads/2019/09/HMS_national-2018-19.pdf