Ha Nguyen

Final Report: School District Closure in the Time of COVID-19, a Social Media Perspective

Abstract

The rapid unfolding of the COVID-19 pandemic in the United States called for immediate response from school districts. While most districts were closed for in-person instruction in March 2020, some districts later transitioned to in-person or hybrid approaches, while others stayed in distance learning. Capturing these varied responses presents a challenge when conducted at scale. In this paper, I present preliminary results of using neural network models to predict when districts were first closed, based on districts' Facebook posts. Results illustrate the promise of these models to understand educational institutions' responses to the emergencies of the pandemic and the utility of using publicly available data and algorithmic approaches for understanding these responses.

Introduction

The use of social media as a platform for educational institution's communication has been on the rise (Kimmons et al., 2019). As an agile tool to keep the public informed (Lovari & Parisi, 2015), social media has been increasingly adopted during the COVID-19 pandemic to disseminate timely responses from public services organizations (Kimmons et al., in press). Inquiring about district responses with district representatives through a survey or an interview can be informative, but is likely expensive to do at scale and may not capture the exact details and timing of the communication. In this paper, I take a public data mining approach: using publicly available, large-scale data for research (Kimmons & Veletsianos, 2018). This approach provides a glance at how responses unfold immediately before and after school closure.

The current study presents preliminary steps to examining the feasibility of such an approach. I explore the following question: **How can we use algorithmic, public data mining approaches on social media posts to understand district communication around school closure?** I leverage multiple neural network models, namely artificial neural network, long-short term memory (LSTM), and convolutional neural network (CNN), in a supervised task to predict the closing dates of school districts in the United States.

With this study, I demonstrate use of data mining approaches with publicly available, rich datasets such as Facebook posts. I detail ways to validate the models, which is an important step to consider as researchers seek to integrate these approaches into educational research. These analyses open up future research opportunities; for example, to extend the analyses to predicting district reopening, or to examine differences in the frequencies and content of districts' communication in case of an emergency, in relation to districts' demographics and student learning following the pandemic.

Background

From 2005 to present, the number of social media posts on Facebook and Twitter from school and school districts has been increasing every year (Kimmons et al., in press). The large existing user base of social media (e.g., 69% of adults in the U.S., approximately 44% of schools; Kimmons et al., 2019; Anderson & Perrin, 2017) makes it a promising venue for educational organizations to communicate with parents and students, among other audience, quickly. Districts can potentially reach parts of the population that may be traditionally underrepresented using social media (Bertot et al., 2012).

Increasing adoption of social media by public agencies is especially the case during emergencies. For example, prior work on government communication has detected spikes in social media communication during critical events, such as earthquake, protests, or public health crises (Mori et al., 2020). Social media can serve two main purposes: as information disseminator or as recipient of feedback (Lindsay, 2011). Social media can be used for community outreach to issue warnings and emergency communications. They can also be used to request information to increase situational awareness and understanding of how agencies can better respond to users' needs (Wukich, 2015).

Researchers on public communication have suggested that organizations may adopt different communication strategies about critical events when the events are one-off versus continuous (Mori et al., 2020). Communication during an ongoing event such as the COVID-19 pandemic will likely ebb and flow as organizations' responses fall into routines (Mori et al., 2020). For example, at the start of the pandemic (March 2020), Kimmons et al. (in press) find a surge in the number of Facebook posts on school district social media channels across the U.S. This surge suggests that districts may have tried to communicate immediate information about school closure and transition to distance learning to parents and students, as communication from the government and public health organizations was unfolding in real time. In my analyses, I observed a similar phenomenon where communication frequency for district Facebook pages slowed down in the months leading to the summer, and picked up again as the Fall semester came around (Kimmons et al.).

While Kimmons et al.'s analyses can reveal the general trends in districts' social media use, such analyses do not illuminate the finer-grained details about how districts disseminate important information regarding school closure, reopening, and resource dissemination in real time. This dissemination may vary with districts' locations (e.g., in areas with a high versus low number of COVID cases), state policies, resource availability, and individual district's decision-making. For example, how promptly do districts provide information about school closure and reopening? How consistent do districts follow up with updates about the pandemic following the initial closure?

Answering these questions would require a large-scale dataset that details districts' communication over time. In this study, I used posts from school districts' social media, a data source that has been largely underexplored in educational research (Kimmons et al., 2019). As a first step, the current paper demonstrates use of different neural network models to predict when a school district was initially closed.

Method

Data Sources

Data sources for this project came from three main sources: Facebook district posts, data from the National Center for Education Statistics (NCES; NCES, 2020) to identify the districts, and a labeled data set of the closing dates of 285 California school districts.

NCES data. The data from this step came from a larger initiative to analyze district social media posts (Rosenberg & Nguyen, in press). The research team first retrieved a list of U.S. school district websites from the publicly available NCES data, along with key demographic variables about the districts (e.g., number of English Language learners, student to teacher ratio, number of students on free and reduced lunch) and the school district websites. From these websites, the team identified all district Facebook pages (n pages = 8,193). These data are all linked to a unique ID for each district (i.e., NCES id).

Facebook district posts. Through Facebook's platform, CrowdTangle (CrowdTangle Team, 2020), the research team retrieved the content of the district' Facebook pages (n matched pages = 7,631), resulting in 3,377,147 posts for 2020. The data from CrowdTangle was then linked to the NCES data via district's Facebook pages.

Labeled data. A next task was to identify a dataset that contained the dates that the districts were closed, to serve as a training set. I retrieved data about when all California school districts was initially closed in March (Edsource, 2020). I then linked this new data source with data from the previous steps using the district names.

In total, the final data set that was used for training and testing consisted of 14,248 posts from 285 California districts. I restricted the publication month of the posts to March to reduce noise from posts too far before or after the start of the pandemic. I created a dichotomous label for whether school was closed (i.e., the post was published on or after the closing date) or not (i.e., before closing date). For example, if the labeled data indicates that a district was closed on March 16, a post would be coded as "closed" if it was published on or after the 16^{th}.

Data Processing & Exploration

The post content was further divided into training (70% of the data) and test sets (30%). The proportion of the classes (i.e., open versus close) was balanced across the training and test sets (45% posts on or after closing for train set; 52% for test set). On average, each post consists of 48.51 words, with a standard deviation of 93.11.

Word Embedding. The data were tokenized prior to modeling using Keras preprocessing. The maximum features were selected at 500. The text was padded or truncated to a maximum length of 100. The text was then transformed into an embedding matrix, where each row number indicated the index of a word, and each column contained the word embedding.

Algorithmic Approaches

In this study, I ran a number of neural network models. A neural net, which is analogously modeled after the human brain, consists of thousands of interconnected nodes organized into layers (Wang, 2003). For example, in a feedforward neural net where information travels upward, a node will assign weights to each incoming input from the nodes below it using an activation function, and multiply the weights with the input data. If the product exceeds a threshold, the input value is passed to the next node. The activation functions help the network learn nonlinear relationships. For example, sigmoid function, which bounds output values between 0 and 1, is common in binary classification tasks. The ReLU function is based on a derivative function to enable evaluation of derivatives between layers from left to right.

A challenge to using feedforward neural net is that it does not account for sequential information from input data. Recurrent Neural Network (RNN) addresses this problem by introducing a recurrent connection (i.e., a loop) on the hidden layer between input and output. A recurrent neuron combines the state of a prior input (e.g., at time t-1) with the current input (e.g., time t) to preserve the relationship between the current and prior inputs. An example case of

RNN is Long Short Term Memory networks (LSTM), which include multiple layers in the recurrent module to store long-term information (Hochreiter & Schmidhuber, 1997).

Finally, I experimented with convolutional neural networks (CNN) to learn the spatial structure of the input, such as spatial arrangement in pictures or word sequence in text. CNNs, which use a sliding window of filters over the input to identify key features, have also shown promise in text classification tasks (Kim, 2014).

I used the word embedding matrix as the feature. Two feature sets were constructed: one where the data was shuffled; and one without shuffling to account for the temporal sequence of the Facebook post. I constructed the following neural network models, each with batch size of 20 and epochs of 10, for each feature set. All models included a last dense layer with sigmoid activation function, and were compiled with a binary cross entropy loss and optimization for accuracy. The models are as follows.

- Model 1 (base model) is a neural network model with the word embedding, a dense layer (with ReLU activation) and a dense layer (with sigmoid function).
- Model 2 (LSTM) applies a LSTM layer following the word embedding layer, setting dropout at 0.2.
- Model 3 (CNN+LSTM) includes a CNN layer, activated with ReLU function, between the word embedding and the LSTM. I used grid search to select the best values for the parameters, and chose filter number of 32, kernel of size 3 for the CNN layer.
- Model 4 (CNN) applies a CNN layer, activated with ReLU function, following the word embedding. The parameters of the CNN are the same as model 3.
- Model 5 (CNN and stacked LSTM) is similar to model 3, but includes another LSTM layer to account for the fact that the posts are ordered by time. The input for the first

LSTM layer is reshaped to correspond to 25 time-steps (i.e., 25 subsequent posts), before passing to the layer with 1 time-step (i.e., current post) as in model 3.

Evaluation

All models were compared across the following metrics: classification accuracy (4275 posts in the test set), accuracy at the date level (31 days in March), accuracy of detecting the date of closing (67 school districts in the test set), and accuracy of predicting reopening (on a purposefully selected districts). Classification accuracy is the rate of correct classifications (i.e., opened or closed). Accuracy at the date level was based on a majority vote—a day was predicted as closed if the models predicted more than 50% of posts on that day as closed.

For detecting the date of closing, I compared the closing dates (from the labeled data) and the earliest date for each district ID that was (1) predicted as "closed" by the models, and (2) followed by four consecutive dates of posting that contained a "closed" prediction. The cutoff of five consecutive dates was based on inspection of data availability: most districts were closed towards the end of the month and did not post every day following the closure. If the dates were similar or off by one day, I coded the accuracy of closing date detection as 1; and if difference was more than one, I coded the accuracy as 0.

I conducted further evaluation with three different types of models that achieved promising results in accuracy and/or predicting the closing dates (baseline, LSTM, CNN and stacked LSTM). In addition to the test set, I created a validation data set from 100 school districts spread across the U.S. This data set consisted of posts from school districts other than California (presenting 25 states in total, with a mix of public and charter districts). The data were manually coded by two undergraduate research assistants for closing and reopening dates. The purpose of this step was to validate the models that were trained on California school districts with other states and thus illuminate the promise of the models in detecting closing dates at scale.

Finally, I explored whether the models not only predicted closing in March, but also reopening in the Fall, even though such data were not present in the training data. Toward this end, I purposefully selected an example school district from the evaluation data sets. I used the model to predict the closing status of posts from the Alpine school district (Utah), which was closed on March 16, 2020 but transitioned to hybrid learning with in-person course in August 2020. I examined whether the models (1) predicted the date of first closing correctly (i.e., a long series of "closed" predictions), and (2) predicted the date of reopening (i.e., marked by a series of "open" predictions).

Fairness

To illustrate a potential approach to exploring model fairness, I evaluated the extent to which there was a difference in accuracy, sensitivity, and specificity between districts based on their demographic characteristics for the three types of models (base models, LSTM, CNN and stacked LSTM). I chose the student to teacher ratio because it likely reflects teacher workload and therefore can serve as a proxy of teacher workload, allocation of resources, and education quality (Card & Krueger, 1992).

I created a dichotomous variable for districts whose student to teacher ratio was in the top quartile (coded as 1 when the ratio equals to or larger than 16.2, which is the mean of the validation data set), and those whose ratio is lower (coded as 0). I fitted models trained on the original training data for subsets of the validation data (student to teacher ratio in the top versus bottom quantile) and compared their accuracy, precision, recall, and f1-score. If there existed a difference in these metrics between the subsets, we may argue that the models may be more or less accurate for certain district subgroups.

Findings

Model Comparisons

Table 1. Results from neural network models

Model	Method	Activation	Accuracy test set (sequential test set)	Accuracy test set (shuffled test set)	Accuracy for date (majority vote)
Sequential (trained on non-shuffled data)					(sequential test set)
Model 1	baseline	sigmoid	.657	.650	.700
Model 2	LSTM	sigmoid	**.694**	.666	.748
Model 3	CNN+LSTM	ReLU, sigmoid	.688	.678	.747
Model 4	CNN	ReLU, sigmoid	.682	**.697**	.754
Model 5	CNN+LSTM+LSTM	ReLU, sigmoid	.683	.667	.732
Shuffled data (trained on shuffled data)					(shuffled test set)
Model 1s	baseline	sigmoid	.663	.667	**.832**
Model 2s	LSTM	sigmoid	.668	.659	.733
Model 3s	CNN+LSTM	ReLU, sigmoid	.690	.652	.741
Model 4s	CNN	ReLU, sigmoid	.689	.661	.742
Model 5s	CNN+LSTM+LSTM	ReLU, sigmoid	.689	.593	.641

Notes. Bolded numbers = highest values.

Table 1 presents results from the neural network models. For the sequential data, models with the LSTM layer (model 2, 3, 5) performed slightly better over 10 epochs. The LSTM model appeared to be the best performing (accuracy of test set at .694). Meanwhile, for the shuffled data, the CNN model performed the best, although examination of the validation loss for this model suggested sign of overfitting (i.e., validation loss for test set increased over time).

Next, I examined the accuracy at the date level (instead of the post level), based on a majority vote. Overall, the models appeared to achieve higher accuracy for this metrics (range of .641 to .832, compared to range of .657 to .694), with the baseline model performing the best (accuracy of majority vote = .832 for the shuffled data).

Regarding the accuracy of detecting the closing dates (Table 2), model 5 and model 1s appeared to have higher accuracy (.612 and .662, respectively). Through manual inspection of the predictions, I did not observe that the LSTM layer had more consistent predictions (i.e., a series of "closed" predictions, as opposed to switching between "closed" and "opened"). In fact, all models seemed to predict closing and opening intermittently rather than sequentially. The accuracy of detecting closing date could be improved for most of the models. I discuss strategies to address the model accuracy in the Discussion.

Table 2. Accuracy of detecting closing date

Model	Method	Activation	Accuracy detecting close date
Model 1	baseline	sigmoid	.478
Model 2	LSTM	sigmoid	.478
Model 3	CNN+LSTM	ReLU, sigmoid	.522
Model 4	CNN	ReLU, sigmoid	.493
Model 5	CNN+LSTM+LSTM	ReLU, sigmoid	**.612**
Model 1s	baseline	sigmoid	**.662**
Model 2s	LSTM	sigmoid	.500
Model 3s	CNN+LSTM	ReLU, sigmoid	.426
Model 4s	CNN	ReLU, sigmoid	.500
Model 5s	CNN+LSTM+LSTM	ReLU, sigmoid	.500

Notes. Bolded numbers = highest values.

Evaluation

I evaluated three types of models in terms of accuracy (model 2; LSTM with sequential data; model 5; CNN+LSTM+LSTM with sequential data; and model 1s; baseline with sequential and shuffled data) against a new data set of districts other than California. At first glance, the models performed quite well (Table 3). However, a closer look at the graphs for accuracy/validation loss over 10 epochs revealed patterns of overfitting for the two base models and the CNN and stacked LSTM (Appendix A). We can observe that the accuracy for the train data was increasing at a much faster rate than the validation data. Similarly, the training loss was decreasing over time for these models, but the validation loss was increasing. Overall, the LSTM model appeared to be the best performing (without overfitting), with the potential to train for a higher number of epochs to increase accuracy.

Table 3. Accuracy metrics for the validation data

Method	Accuracy	Precision	Recall	f1-score
Baseline (with shuffled data)	.756	.756	.758	.758
Baseline (with sequential data)	.718	.768	.750	.759
LSTM	.746	.750	.821	.784
CNN+LSTM	.764	.783	.802	.793

Next, I explored use of the three models to predict when a school district reopened. Table 4 presents example posts and predictions from the three models. The models all picked up the posts related to COVID-19. However, they were not able to predict the "reopening" date (August 18), and rather switched back and forth between "closed" and "opened" in the summer months when the districts were closed. This pattern was not surprising given that the models likely picked up words specific to COVID-19 in the days after closing, so any post contained words such as "updates about COVID-19", "opening schedule", or "connectivity issue" would likely be classified as "closed", whereas the district had these posts even after reopening.

Table 4. Predictions of "closing" and "opening" from the best performing models for a sample

Message	Date	Real status	Base	LSTM	CNN+ LSTM+ LSTM
Due to the emerging concerns associated with the COVID-19 virus, the Board of Education and District Administration have adjusted the school schedule for Monday and Tuesday, March 16-17. The schedule for these two days will be a minimal day schedule as outlined on our webpage…	3/12	open	open	closed	closed
Here's the message from our webpage: alpineschools.org	3/12	open	open	open	open
Dear Alpine School District Parents, Following today's press conference with Governor Herbert and in an attempt to proactively respond to the COVID-19 pandemic, all Utah schools have been dismissed for two weeks beginning Monday, March 16, 2020. We have provided more information on our alpineschools.org webpage...	3/13	open	closed	closed	closed
We love stories like these! Here is another one of our amazing teachers spreading #ASDshine.	3/13	open	open	closed	open
SHINE ALPINE! Thanks to all of our employees who are currently working to get online learning platforms ready to launch for tomorrow…	3/16	open	open	open	open
For more information go to: https://health.utahcounty.gov/2020/03/18/covid19publichealthorder/	3/17	open	closed	closed	closed
With final numbers still being gathered, we estimate around 10,000 Chromebooks were checked out today. Thanks to dedicated staff who worked	3/17	open	open	open	closed

to prepare devices and check them out in an orderly and safe manner. *CHROMEBOOK CONNECTIVITY* …					
Wednesday's Meals Served: 1,778 breakfasts served (34% of our daily average) 3,715 lunches served (11% of our daily average) #ASDShine	3/17	open	closed	closed	closed
ASD CHROMEBOOK CONNECTIVITY: As many of you may have noticed, there is significant delay in connectivity for online learning at this point. Our Chromebooks are slow to connect or are not connecting at all. Alpine is working with the state and vendors to remedy the issue asap.	3/18	closed	closed	closed	closed
Stay calm and carry on! Parents, we continue to be amazed by your efforts to organize your homes as remote learning sites. Thank you for partnering with us. ASD parents are a big part of our Shine! #ASDshine	3/18	closed	closed	closed	closed
Notice from Utah Department of Natural Resources: "The rumor of an imminent larger earthquake is incorrect. While anything is possible, it's unlikely. Our experts, along with experts from USGS, indicate the probability of another 5.0 magnitude earthquake in the next week is relatively low. #utquake #utahearthquake	3/18	closed	closed	closed	closed
Computer device checkout underway at all of our schools. Thank you for helping to make today happen! Staying safe and getting this done as quickly and organized as possible.	3/18	closed	closed	closed	closed
[**District stayed closed until transitioning to hybrid**]					
Good Day! Go to alpineschools.org for updated information on Return to Learn plans as well as a Q & A section. Here is an example of a Question: Is ok for students to wear a face shield instead of a mask? Answer: Yes…	7/16	closed	closed	closed	closed
4th Annual Tools For Schools sponsored by Jordan Credit Union will be held at Shops At South Town (10450 South State Street in Sandy, south of the mall in the Macy,Äôs parking lot) starting Tuesday, August 11 at 3:00 pm until Thursday, August 13 at 7:00 pm. We are mindful of COVID-19 restrictions and there is no need to get out of your vehicle. We will have a convenient drive and drop donation area…	8/11	closed	open	closed	open
We are excited to get things off to a great start this morning! Welcome back to school, everyone! #alpineschools #onecommunityonevision	**8/18**	**open**	open	closed	open
Welcome Back Alpine! It has been a busy week, and we appreciate everyone's work, cooperation, and patience as the new school year has started. Have a safe and restful weekend! #ONEAlpine Alpine School District #Alpineschools	**8/21**	**open**	open	closed	open
Congratulations to Greenwood Gators on the ribbon cutting for their school rebuild! Thanks to everyone who made this project possible. Go Gators!! @Alpine School District #alpineschools	**8/26**	**open**	open	open	open
Attention On-line Students, Parents, and Teachers...We are experiencing district-wide network issues this morning and are working to resolve the issues as quickly as possible. Thanks for your patience!	**8/26**	**open**	closed	closed	closed

Fairness

Across the models (baseline, LSTM, CNN and stacked LSTM), I examined difference between the district subgroups (top and bottom quantiles in terms of student to teacher ratio) in accuracy, precision, recall, and f1 score. The results (Table 5) are quite similar to metrics from models conducted with the entire data set (Table 3). The majority of the models (except for LSTM) had higher recall and f1-score for districts in the top quantile of student to teacher ratio. A low recall rate may suggest overfitting and/or unbalanced classes. Because the two classes in the prediction (opened or closed) were quite balanced in the training set, future research can consider further tuning of regularization parameters.

In production, usage of the models, particularly the baseline models, warrants more examination of the recall rates across the district subgroups. This is because the recall rates for the baseline models for the bottom quantile were low (.68), compared to the top quantile of districts (.84). Future work can also examine fairness in the model prediction, by including and excluding other variables in the predictions and comparing accuracy, such as those with higher versus low number of posting, those with high versus low poverty, or those with high versus low number of COVID cases.

Table 5. Accuracy metrics for district subgroups

	Top quantile (>= mean ratio)				Bottom quantile (< mean ratio)			
Method	Accuracy	Precision	Recall	f1	Accuracy	Precision	Recall	f1
Baseline (shuffled data)	.770	.755	.844	.797	.758	.758	.680	.717
Baseline (sequential data)	.755	.755	.844	.797	.758	.758	.680	.717
LSTM	.720	.774	.743	.758	.742	.725	.835	.776
CNN+LSTM	.785	.771	.850	.809	.743	.796	.758	.777

Discussion

Examining how districts communicate with parents, students, and the larger public during an unprecedented emergency such as COVID-19 is challenging at scale. In this study, I used neural network models to predict when a district was closed or opened based on district Facebook posts. This study serves as the first step to illustrate the potential of these approaches. I found that models trained on just the text from Facebook posts could predict whether the district was closed or opened in a supervised task within the first month of closing, although future work remains to improve the accuracy for detecting closing and reopening dates reliably.

The analyses can be extended to longer time periods (e.g., prior to and after the pandemic has started), since understanding districts' communication can illuminate how the pandemic has altered educational practices. Accurately documenting when districts opened or closed reflects the varied responses to the pandemic, inviting future research into how these responses may vary with the emergencies of the pandemic in different locations, the demographics that the districts serve, and their underlying political beliefs. Such information is valuable, because to the best of my knowledge, there is no publicly available, systematic documentation of closing-reopening schedules for school districts in the United States.

Results from the analyses call for a reexamination of how we define accuracy. Most models did not perform well when predicting the closing dates (within a one day off accuracy). In future iterations, I will experiment with varying the accuracy thresholds. For example, an idea is to adjust the threshold from one day to one week of the closing date.

Another challenge is the overfitting observed in the baseline and models with stacked LSTM layers. Future steps include increasing the amount of data for the training test through data collection and augmentation (e.g., scaling, rotation), introducing multivariate data (e.g., types of external links in the posts, whether posts included images or not, etc.), simplifying the models, introducing early stopping, or introducing higher dropouts in the CNN and LSTM layers. Finally, to improve the model accuracy, I will experiment with other word embedding models that have shown state-of-the-art results in classification tasks.

Conclusion

This study has implications beyond the COVID-19 pandemic. Given the local nature of education in the U.S., social media remains a pertinent channel where districts communicate with stakeholders such as parents and students. However, social media posts by districts have not been widely examined in educational data mining. I demonstrate the potential of using social media posts by educational institutions to examine local decision-making at scale. Findings also illuminate the challenges in finding appropriate accuracy metrics for model evaluation.

References

Anderson, M., & Perrin, A. (2017). Technology use among seniors. *Washington, DC: Pew Research Center for Internet & Technology*.

Bertot, J. C., Jaeger, P. T., & Hansen, D. (2012). The impact of polices on government social media usage: Issues, challenges, and recommendations. *Government information quarterly*, *29*(1), 30-40.

Card, D., & Krueger, A. B. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of political Economy*, *100*(1), 1-40.

CrowdTangle Team (2020). CrowdTangle. Facebook, Menlo Park, California, United States. List ID: all-k12-institutions

Johnson, S. (2020). List of California K-12 districts closed for in-person instruction due to coronavirus. https://edsource.org/2020/california-k-12-schools-closed-due-to-the-coronavirus/624984

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

Kim, Y. (2014). Convolutional neural networks for sentence classification. In *Proceedings of Empirical Methods on Natural Language Processing*.

Kimmons, R., Hunsaker, E. W., Jones, J. E., & Stauffer, M. (2019). The Nationwide Landscape of K–12 School Websites in the United States. *The International Review of Research in Open and Distributed Learning*, *20*(3).

Kimmons, R., & Veletsianos, G. (2018). Public internet data mining methods in instructional design, educational technology, and online learning research. TechTrends, 62(5), 492-500. doi:10.1007/s11528-018-0307-4

Kimmons, Rosenberg, & Allman (in press). Trends in Educational Technology: What Facebook, Twitter, and Scopus can Tell Us about Current Research and Practice. *EdTech Trends*.

Lindsay, B. R. (2011). Social media and disasters: Current uses, future options, and policy considerations.

Lovari, A., & Parisi, L. (2015). Listening to digital publics. Investigating citizens' voices and engagement within Italian municipalities' Facebook Pages. *Public relations review*, *41*(2), 205-213.

Mori, E., Barabaschi, B., Cantoni, F., & Virtuani, R. (2020). Local governments' communication through Facebook. Evidences from COVID-19 pandemic in Italy. *Journal of public affairs*, e2551.

National Center for Education Statistics. (2020). ELSI Table Generator. https://nces.ed.gov/ccd/elsi/tableGenerator.aspx

Rosenberg, J., & Nguyen, H. (in press). How K-12 School Districts Communicated During the COVID-19 Pandemic: A Study Using Facebook Data. Poster presented at The 11th International Conference on Learning Analytics and Knowledge (LAK21)

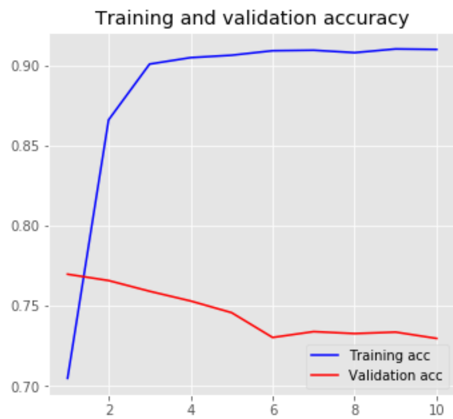
Wang, S. C. (2003). Artificial neural network. In *Interdisciplinary computing in java programming* (pp. 81-100). Springer, Boston, MA.

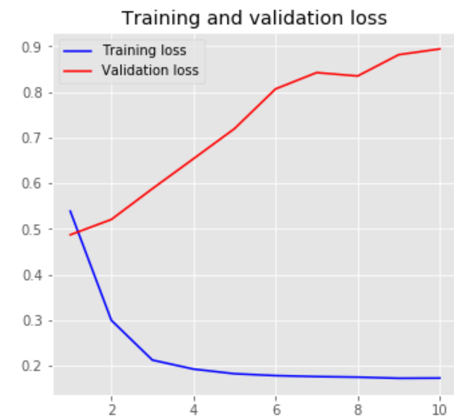
Wukich, C. (2015). Social media use in emergency management. *Journal of Emergency Management*, *13*(4), 281-294.

Appendix

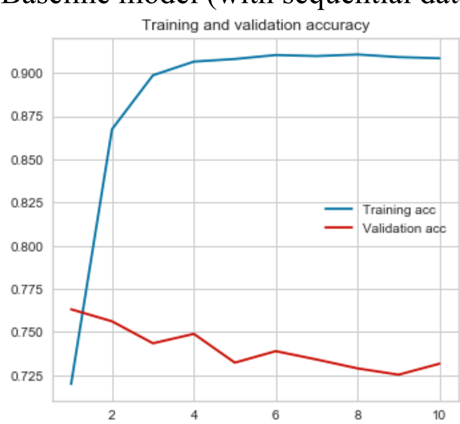
Training and validation loss for the 3 best performing models on the validation sets

A. Baseline model (with shuffled data)

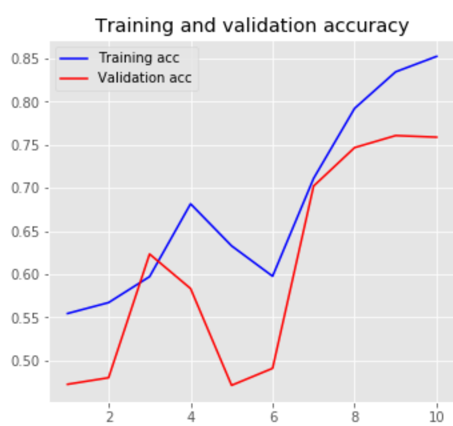


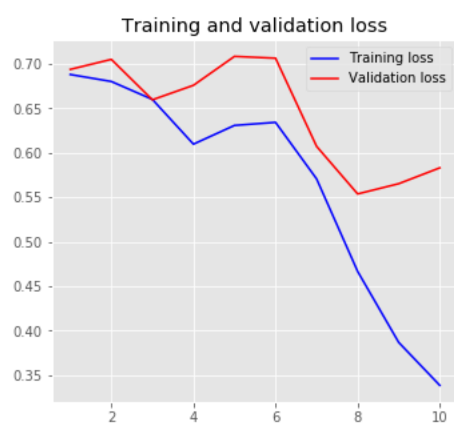


Baseline model (with sequential data)

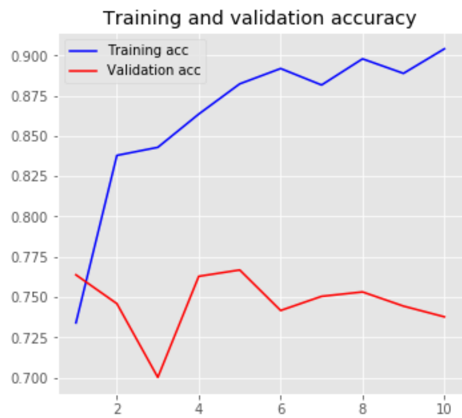


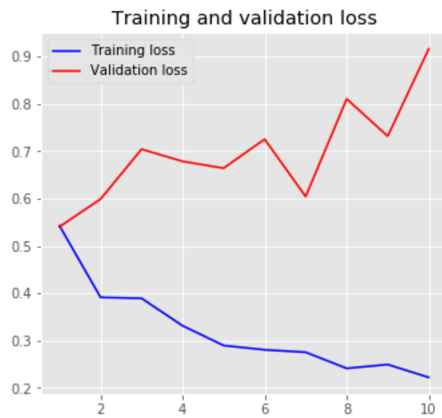
B. LSTM





C. CNN+LSTM+LSTM


Training and validation accuracy
Training acc
Validation acc
0.900
0.875
0.850
0.825
0.800
0.775
0.750
0.725
0.700
2
4
6
8
10


Training and validation loss
Training loss
Validation loss
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
2
4
6
8
10