

Exploring Your Areas (Spring 2019) Assignment 3

Vincent La*

January 26, 2019

Abstract

This assignment narrows in on a few topic areas of focus. We will look at a variety of types of sources to survey how the field has targeted different audiences, different specific content areas, and different technologies. We will focus our discussion around seven different categories: Technology, Audience, Content, Sociotechnical Issue, Theory, Medium, and Methodology. We will focus mainly on three Topic Areas: Applications of Natural Language Processing in Education (with a particular focus on detection of fake news), how to evaluate effectiveness of teachers, and the effect of education on socioeconomic outcomes (like housing prices, health, etc.)

*vincent.la@gatech.edu

1 Topic 1: Applications of Natural Language Processing in Education (with a particular focus on detect of Fake news)

The topic that I've come to be most passionate about is the application of Natural Language Processing in Education. This topic will very likely be the topic that I pursue for final project. In particular, I've been particularly passionate about trying to tackle the issue concerning Fake News.

To date, there have been many applications of Deep Learning and Natural Language Processing in the Education space. For example, Ferreira (2018) [12], discussed in the International Conference on Artificial Intelligence in Education in 2018, describes how to use Natural Language Processing and Topic Modeling to evaluate students' use of asynchronous discussions (e.g. on an Online Platform like Piazza) in online learning environments. What's interesting is how the paper examines students' cognitive development across different course topics. Burstein (2007) [8] also goes through a survey of how NLP has been applied to Education related topics.

To me, one very interesting usage of NLP and deep learning is around the identification of "Fake News". As Bovet (2019) [6] examine, the influence and scale of Fake news exponentially escalated in 2016 with the influx of mass-scale Social Media over the past decade. In Bovet (2019) [6] the authors investigate the influence of fake news in Twitter during the 2016 Presidential Election. Using a data set of 171 million tweets in the five months preceding election day, 25% of these tweets spread either fake or extremely biased news. In addition, Polletta (2019) [25] provide deep stories and narratives around storytelling and the role of fake news in the Trump era. Furthermore, Next, Allcott [2] discusses the usage of fake news in the 2016 election circulated through Social Media. I find this a very core *Education* issue as Political Education, or the general public's knowledge about the true state of our government and politics is very important. In addition, Pennycook (2018) [24] investigates the type of people that are likely to believe in fake news. Given that fake news is extremely easy to disseminate and relatively much harder to disprove, the potential negative impact of fake news on our public's education and trust of the system is immense, and solving this problem is one of the biggest and truly unprecedented problems of the 21st century.

On existing literature, there has been many papers that have examined trying to build machine learning to detect fake news. Gurav (2019) [19] provides a survey study on Automated Systems developed for Fake News Detection using NLP & Other Machine Learning Techniques. In this paper, they cite Gilda (2017) [16] who shows how NLP can be relevant to detect fake information. They used time period frequency-inverse frequency (TF-IDF) of bi-grams and probabilistic context free grammar detection. Buntain (2017) [7] use Twitter data and use NLP and ML to predict Twitter tweets that are propagating fake news. To do this, they use two-credibility focused Twitter datasets - CREDBANK, a crowdsourced dataset of accuracy assessments for events in Twitter, and PHEME, a dataset of potential rumors in Twitter and journalistic assessments of their accuracies. They apply this model to

Twitter content sourced from BuzzFeed’s fake news data set and show that models trained against crowdsourced workers outperform models based on journalists’ assessments and models trained on a pooled dataset of both crowdsourced workers and journalists. Rashkin (2017) [27] uses NLP to find the types of language and wording that is typically used in fake news. Conroy (2015) [9] also discuss methodologies used to detect fake news.

In addition to research papers that use existing datasets and conduct NLP and Machine Learning to predict Fake news, Diana (2018) [11] proposes one step further. This paper was initially presented in the Artificial Intelligence in Education 2018 Conference. This paper offers an interesting perspective: to use AI to build a process to classify what is “Fake news” vs real. This potentially gives a chance to classify such news at scale. To be more specific, in this paper, they propose building a tool that allows users to become educated on what fake news is and how one would detect fake news. This is important as models are not necessarily sufficient. We need to educate people on how to use their own judgement and figure out what is fake vs what is real. Second, the paper proposes building an app where users could label sources as fake or not. This is important as this allows the construction of bigger and bigger labeled datasets which could then be used for Machine Learning and improving our algorithms.

In terms of industry applications, Eyeo GmbH (Wladimir Palant), a German software development company, has produced a Fake News Detector as a Google Chrome Extension: <https://chrome.google.com/webstore/detail/trusted-news-for-google-c/koejmcafidkcjlnckgkpfjfbijkhkpchei> (Trusted News for Google Chrome). There are similar tools as well, for example Fake News Detector which is open sourced: <https://github.com/fake-news-detector>.

All of these tools offer interesting solutions and machine learning models that can be applied to what is in the user’s browsers. However, there are certainly additions that can be made and improvements. For example, none of these tools allow users to add to the central datasets and add labeled data to allow for iteration and feedback to improve the machine learning models. For example, I think a great addition would be a combination of papers and industry tools cited above in addition to Wang (2017) [31] to contribute back to baseline datasets that can be used for further Machine Learning.

2 Topic 2: Evaluating Teachers: The important role of value-added

The next topic that I discuss is on methodologies to evaluate Teachers, primarily K-12 teachers; in particular the role of “value-added” measures. The call to reform how we evaluate teachers in the USA has been for many decades now. For example, Haefele (1993) [20] discuss a “Call for Change” in how we evaluate teachers. In this paper, Haefele discusses how the “norm” of teacher evaluation is very anecdotal and subjectively. In the vast majority of school districts, a non-tenured teacher is observed twice a year, for a period of 20 to 30 minutes, by the principal. Tenured teachers are observed for approximately the same num-

ber of minutes, but less frequently. Sometimes a conference follows the observation where the teacher receives feedback; sometimes this does not even happen. Of course, this method of evaluation is troubled. Evaluation criteria lack objectivity and validity. Evaluators are untrained. However, attempts to overthrow this method in American schooling appears unachievable. In the 2.5 decades since Haefele (1993), new methods have been developed, “high stakes testing” has become a reality for many school districts; yet, the answer of how we actually evaluate teachers optimally remains elusive.

This topic is important to me (although admittedly not as important to me as detection of fake news) because evaluation of teachers and in general public education reform is critical. Education is often considered the “great equalizer” in socioeconomic equality, but if students are placed into a bad classroom they stand a worse chance at achieving better outcomes than students who are placed in a good classroom.

High stakes testing really started in the 1980’s to 1990’s, but didn’t become widespread popular until the early 2000’s, as Amrein (2002) [3] discuss. In 1983, the National Commission on Education released *A Nation at Risk* [23]. *A Nation at Risk* called for an end to minimum competency testing movement and the beginning of a high-stakes testing movement that would raise the nation’s standards of achievement drastically. The Commission recommended that states institute high standards to homogenize and improve curricula and rigorous assessments be conducted to hold schools accountable for meeting those standards. The effects of this report affect America still today.

Perhaps the most infamous case in recent history was the Washington DC School District in the early 2010’s. As Klein (2010) [13] discuss, Joel Klein (then Chancellor of New York City Department of Education) and Michelle Rhee (then chancellor, of District of Columbia Public Schools) released a manifesto on how to “save” America’s schools. They emphasized that the key to education is on focusing on the quality of the teacher and to do away with teacher hiring and retention standards that for far too long purely involved seniority and academic credentials. They focused more on evaluating teachers, and promoting teachers based on performance, and even removing teachers who were underperforming, regardless of seniority.

Of course, the problem is, how do we evaluate teachers? In person review, as discussed earlier, is a start but has many faults, including not scalable and very labor intensive, and unclear guidelines and untrained evaluators. Thus came the movement of “Value-Added” measures with high-stakes testing. The seminal paper on applying “Value-Added” methodologies within education is McCaffrey (2004) [22]. The basic premise of this methodology is to use standardized test scores and compare a teacher’s students’ performance on their standardized tests compared to their performance last year normalized for student quality so that the scores for teachers can be compared across other teachers and school districts. The idea here is to separate out and measure the causal effect or the “value-added” of a particular teacher on a student’s learning outcomes.

One source that talks about this is Glazerman (2010) [17]. This paper was published in

the Brown Center on Education Policy at Brookings. This source is a survey study on “Value-Added” methodologies to measure effectiveness of teachers (generally in the K-12 range). In summary, “Value-Added” refers to the measurement of the evaluation of teachers based on the contribution they make to the learning of their students. However, it has been controversial since many Value-Added methodologies use standardized test scores as the base data. Standardized test scores may not fully capture the nuance of the true contributions that a teacher makes on a student. Furthermore, it is very difficult to separate out value added of the teacher vs student characteristics. In particular, naively looking at test scores of schools in worse neighborhoods correlated with worse socioeconomic characteristics would imply systems level problems, not teacher-specific.

That being said, this paper discusses the merits of “Value-Added” methodologies. Current methods are clearly not working. In many school districts where the teachers are evaluated on a “binary” metric (e.g. Satisfactory vs not) most teachers are given “satisfactory” evaluations with no differentiation between teachers who contribute a lot and are high performing vs low-performing teachers. In addition, “Value-Added” methodologies, although imperfect have been shown to have definitely non-zero relationship with educational outcomes.

For a potential research project, it may be interesting to get standardized test scores from a public school district (or even Georgia Tech OMSCS program!) and apply Value-added measures to try to rate the quality of instructors. Information could be used to inform policymakers and school administrators on which instructors need better training, or students on which classes to pick. I can imagine that this project could be both on the “Research” track and “Development” track as there’s analysis to do but one can imagine a web app where administrators can plug in student longitudinal record of scores to compute value-added measures.

3 Topic 3: Effect of Education on Socioeconomic Outcomes

The third topic that I explore is the effect of education on Socioeconomic Outcomes. It is often said that education is important for all children; in fact, in America and most other countries, compulsory education is the norm. But what exactly are the effects of education on socioeconomic outcomes, and what are the different types of education out there that affect different outcomes. For example, there is not only traditional school, but trade classes, online courses, seminars, and a countless other forms how people learn. For this topic, I wanted to dive deep on different forms of education and the outcomes that they affect.

For my first source, I cite a research paper Goodman (2019) [18] as it researches whether or not Online Delivery increases Access to Education. Furthermore, it uses data from OMSCS which I thought was really interesting!

The paper was originally distributed as a working paper via National Bureau of Economic

Research and was eventually published in the Journal of Labor Economics. The paper is the first that investigates whether online education actually affects the number of people pursuing education. Using the OMSCS program, the paper uses a Regression Discontinuity approach with exploits an admissions threshold that shows that access to the online option actually does increase enrollment into education. In conclusion, it estimates that OMSCS will boost the annual production of US CS Masters Degrees by about seven percent!

The technology that the source uses is Data Analytics/Statistics primarily using Regression discontinuity technique to try to estimate causal effects of the program on number of people pursuing education. Specifically, in the very first cohort, the program's admission officer read applications in descending order of undergraduate GPA until he had identified about 500 applicants to which immediate admission was offered. As a result, such offers were made only to those with a GPA of at least 3.26, a threshold that was arbitrary and unknown to applicants. The officer eventually read all of the applications and some of those below the threshold were offered deferred admission. A regression discontinuity design shows this admissions process created at the threshold a roughly 20 percentage point difference in the probability of admission to the online program. The paper then used National Student Clearing House data to compare enrollment outcomes for those applications just under vs just over the threshold.

This first paper touched on the effect of an establishment of a new education provider (OMSCS!) on education outcomes itself! But what about a alternative forms of education. For example, Agarwal (2010 [1]) looks at how voluntary financial education can affect loan default rates during a housing crises. The specific housing crisis they examined was the most recent one during the Great Recession of 2008. Using data from the Indianapolis Neighborhood Housing Partnership Data, the authors performed Econometric analysis and found a significant effect that individuals who received loans counseling and education on the terms of their mortgage performed substantially better in terms of loan default outcomes then others who did not complete those education courses. I found this extremely interesting as this is not formal education, but influenced a very important outcome!

Next, I look into the causal effect of education on health. Siles (2009) [28] examines specifically this using data from the United Kingdom. Cutler (2006) [10] also looks into the causal effect of education on health. In Siles (2009), the paper uses changes in compulsory schooling laws in the United Kingdom to test the hypothesis around the link of years of schooling to indicators of health. Several different types of health metrics were used. The results provide evidence of a causal relation running from more schooling to better health which is larger than the current literature suggests!

Especially, in the US, schooling is often tied very specifically to where you live. This concept of a "school catchment zone", or where you live being a direct relationship to where you go to school can often be reflected in parents' buying choices. That is, many parents who care about schooling will choose to locate their house in areas with better schools. Of course, if this results in an increase in demand, it's interesting and relevant for policymakers what the effect of capitalization of school quality into housing prices is.

Black (1999) [4] wrote the seminal paper on this, by using a Regression Discontinuity technique along the border of school catchment areas. While the question seems intuitive, the estimation of the causal effect of non-trivial. Housing prices are positively correlated with school quality, but better schools are often located in areas with better amenities that reflect higher housing prices. Thus, it is very difficult to separate the effect of school quality into housing prices. Black (1999) cleverly estimated this by looking at the boundary of school catchment areas. By looking at the housing prices just across the boundary, she can see what the effect is of the school quality on housing prices since houses across the boundary are probably very similar and have access to the same local amenities.

For a potential research project, it would be interesting to focus on a socioeconomic outcome, such as ones discussed above and get data and explore the effect of education (school quality, years of schooling, etc.) on the socioeconomic outcome. In fact, this is what I did many years ago as a researcher myself! In La (2015) [14], which is a paper I wrote myself in 2015, I looked at the capitalization of school quality into housing prices using Walk Zones from the Boston Public School District. That is, around each school within one mile radius (or walking distance), students living in that area had a higher probability of entering that school than students not living in that area. I found significant effects of housing prices within areas of a good school vs a bad school.

References

- [1] Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., & Evanoff, D. D. (2010). Learning to cope: Voluntary financial education and loan performance during a housing crisis. *American Economic Review*, 100(2), 495-500.
- [2] Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211-36.
- [3] Amrein, A. L., & Berliner, D. C. (2002). High-stakes testing & student learning. *Education policy analysis archives*, 10, 18.
- [4] Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *The Quarterly Journal of Economics*, 114(2), 577-599.
- [5] Bosch, N., Mills, C., Wammes, J. D., & Smilek, D. (2018, June). Quantifying Classroom Instructor Dynamics with Computer Vision. In *International Conference on Artificial Intelligence in Education* (pp. 30-42). Springer, Cham.
- [6] Bovet, A., & Makse, H. A. (2019). Influence of fake news in Twitter during the 2016 US presidential election. *Nature communications*, 10(1), 7.
- [7] Buntain, C., & Golbeck, J. (2017, November). Automatically Identifying Fake News in Popular Twitter Threads. In *Smart Cloud (SmartCloud), 2017 IEEE International Conference on* (pp. 208-215). IEEE.

- [8] Burstein, J. (2009, March). Opportunities for natural language processing research in education. In *International Conference on Intelligent Text Processing and Computational Linguistics* (pp. 6-27). Springer, Berlin, Heidelberg.
- [9] Conroy, N. J., Rubin, V. L., & Chen, Y. (2015, November). Automatic deception detection: Methods for finding fake news. In *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community* (p. 82). American Society for Information Science.
- [10] Cutler, D. M., & Lleras-Muney, A. (2006). Education and health: evaluating theories and evidence (No. w12352). National bureau of economic research.
- [11] Diana, N. (2018, June). Leveraging Educational Technology to Improve the Quality of Civil Discourse. In *International Conference on Artificial Intelligence in Education* (pp. 517-520). Springer, Cham.
- [12] Ferreira, R., Kovanovic, V., Gasevic, D., Rolim, V. (2018, June). Towards Combined Network and Text Analytics of Student Discourse in Online Discussions. In *International Conference on Artificial Intelligence in Education* (pp. 111-126). Springer, Cham.
- [13] Klein, J. (2010). How to fix our schools: A manifesto by Joel Klein, Michelle Rhee and other education leaders. *The Washington Post*, 10.
- [14] La, V. (2015). Capitalization of school quality into housing prices: Evidence from Boston Public School district walk zones. *Economics Letters*, 134, 102-106.
- [15] Le, C. V., Pardos, Z. A., Meyer, S. D., & Thorp, R. (2018, June). Communication at Scale in a MOOC Using Predictive Engagement Analytics. In *International Conference on Artificial Intelligence in Education* (pp. 239-252). Springer, Cham.
- [16] Gilda, S. (2017, December). Evaluating machine learning algorithms for fake news detection. In *Research and Development (SCORED), 2017 IEEE 15th Student Conference on* (pp. 110-115). IEEE.
- [17] Glazerman, S., Loeb, S., Goldhaber, D., Raudenbush, D., Staiger, D., & Whitehurst, G.J. (2010). Evaluating teachers: The important role of value-added. The Brookings Brown Center.
- [18] Goodman, J., Melkers, J., & Pallais, A. (2019). Can online delivery increase access to education?. *Journal of Labor Economics*, 37(1), 000-000.
- [19] Gurav, S., Sase, S., Shinde, S., Wabale, P., & Hirve, S. (2019). Survey on Automated System for Fake News Detection using NLP & Machine Learning Approach.
- [20] Haefele, D. L. (1993). Evaluating teachers: A call for change. *Journal of Personnel Evaluation in Education*, 7(1), 21-31.

- [21] Lubold, N., Walker, E., Pon-Barry, H., & Ogan, A. (2018, June). Automated Pitch Convergence Improves Learning in a Social, Teachable Robot for Middle School Mathematics. In *International Conference on Artificial Intelligence in Education* (pp. 282-296). Springer, Cham.
- [22] McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of educational and behavioral statistics*, 29(1), 67-101.
- [23] National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. *The Elementary School Journal*, 84(2), 113-130.
- [24] Pennycook, G., & Rand, D. G. (2018). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking.
- [25] Polletta, F., & Callahan, J. (2019). Deep stories, nostalgia narratives, and fake news: Storytelling in the Trump era. In *Politics of Meaning/Meaning of Politics* (pp. 55-73). Palgrave Macmillan, Cham.
- [26] Pfeifer, A., & Lugin, B. (2018, June). Female Robots as Role-Models?-The Influence of Robot Gender and Learning Materials on Learning Success. In *International Conference on Artificial Intelligence in Education* (pp. 276-280). Springer, Cham.
- [27] Rashkin, H., Choi, E., Jang, J. Y., Volkova, S., & Choi, Y. (2017). Truth of varying shades: Analyzing language in fake news and political fact-checking. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (pp. 2931-2937).
- [28] Silles, M. A. (2009). The causal effect of education on health: Evidence from the United Kingdom. *Economics of Education review*, 28(1), 122-128.
- [29] Smith, V. C., Lange, A., & Huston, D. R. (2012). Predictive modeling to forecast student outcomes and drive effective interventions in online community college courses. *Journal of Asynchronous Learning Networks*, 16(3), 51-61.
- [30] Tang, C., Ouyang, Y., Rong, W., Zhang, J., & Xiong, Z. (2018, June). Time Series Model for Predicting Dropout in Massive Open Online Courses. In *International Conference on Artificial Intelligence in Education* (pp. 353-357). Springer, Cham.
- [31] Wang, W. Y. (2017). "Liar, liar pants on fire": A new benchmark dataset for fake news detection. *arXiv preprint arXiv:1705.00648*.