JOSEPH LILLEBERG

+1 (507) 829-8598 ⋄ jjglillebergds@gmail.com ⋄ github.com/jlilleberg ⋄ linkedin.com/in/jjglilleberg ⋄ jlilleberg.com

EDUCATION

Master's in Computer Science *Sept. 2016 - June 2018*
University of California, San Diego CGPA : 3.49/4.0

B.S. Computer Science, B.A. Mathematics *Aug. 2012 - May 2016*
Southwest Minnesota State University CGPA : 3.81/4.0

SKILLS

Programming Languages & Tools: Python, R, SQL, JavaScript, Java — Tableau, Git and Version Control
Libraries & Packages: NumPy, Pandas, Matplotlib, Scikit-Learn, Tensorflow, Seaborn, SciPy, Statsmodels, Spacy, Plotly
Technical: Machine Learning, Deep Learning — Neural Networks, Data Wrangling, Webscraping, APIs, Visualization, Decision Analysis, Predictive Modeling, Forecasting, Probability, Statistics, Multivariate Calculus, Linear Algebra

EXPERIENCE

Self-Reflection: Web Development and Machine Learning Marshall, Minneosta
Unemployed *June 2018 - Present*

- **Full Stack (Oct. 2018 - June 2019):** Pursued and applied for Full Stack Development learning JavaScript, React, and NodeJS which culminated in a personal website portfolio.
- **Data Science (June 2019 - Present):** Studied and practiced core data science and mathematical concepts including cleaning messy data, visualizations, machine learning, deep learning, modeling, neural networks, and analysis.

NSF funded research internship in Computer Security Denton, Texas
University of North Texas *June 2015 - Aug. 2015*

- Selected as one of 10 participants nationwide for a 10-month internship researching brain wave scanners in computer security.
- Developed an application that uses changes in EEG wave patterns to detect subconscious recognition. — **C#, Java, Python**

NSF funded research internship in High Performance Data Mining Atlanta, Geogia
Georgia State University *May 2014 - July 2014*

- Selected as one of eight participants nationwide for a 10-month internship researching classification using Google's Word2vec.
- Classified 18,000 documents with 89.73% accuracy using an aggregate model of wod2vec weighted by tf-idf w/o stopwords and tf-idf without stop words. Resulted in a first author publication in IEEE (see below). — **Python**

SELECTED PROJECTS

Topic Modeling, Sentiment and Textual Analysis of U.S. Presidential Transcripts *Dec. 2019 - Mar. 2020*

- Webscraped and cleaned 992 official presidential transcripts, up to Sept. 25th, 2019, consisting of 3.8+ million words or 22+ million characters using Spacy.
- Performed qualitative data analysis, topic modeling, and sentiment analysis by generating 288 interactive visualizations on term associations, empaths, term frequencies, topic frequencies, and word similarities of transcripts for each political era.
- Implemented a RNN to generate artificial transcripts for any given president using TensorFlow. — **Python**

Predicting Medical Costs *Nov. 2019 - Dec. 2019*

- Implemented a XGBRegressor model using scikit-learn's polynomial features, pipelines, feature selection and importance to predict medical costs with an adjusted R^2 of .846, 84.6% goodness of fit. — **Python**

LEADERSHIP AND AWARDS

- **Math & Computer Science Club:** Elected President and Vice President for 3 consecutive years.
- **League of Legends Esports Club:** Founded the LoL club which has been established as an official Esport at SMSU.
- **ACM-International Collegiate Programing Contest:** Honorable mention for two consecutive years.

PUBLICATIONS

- J. Lilleberg, Y. Zhu and Y. Zhang, "Support vector machines and Word2vec for text classification with semantic features," *2015 IEEE 14th International Conference on Cognitive Informatics Cognitive Computing (ICCI*CC)*, Beijing, 2015, pp. 136-140, doi: 10.1109/ICCI-CC.2015.7259377.