Subhankar Ghosh

https://www.linkedin.com/in/subhankar-19 https://subhankar-ghosh.github.io/
ghosh17@illinois.edu Address: 48 E John St, Champaign, IL Mobile : +1-217-904-5807

Education

- **University of Illinois at Urbana-Champaign** GPA: (3.93/4.0)
  Master of Science in Statistics *Aug 2017 – May 2019*
- **National Institute of Technology, Rourkela, India** GPA: (8.95/10.0)
  Bachelor of Technology in Computer Science & Engineering *Aug 2011 – May 2015*

Relevant Coursework

Applied Regression and Design, Statistical Learning/Machine Learning, Neural Networks and Deep Learning, Improving Deep Neural Networks, Convolutional Neural Networks, Sequence Models, Deep Learning for NLP

Computer Skills

- C++, Python, Tensorflow, Keras, NoSQL, SQL, R, Matlab (Limited Experience), Linux

Experience

- **Google** Pittsburgh, Pennsylvania
  Software Engineering Intern *May 2018 - Aug 2018*
  - **Display Data Anomaly Detection**: Detected source outliers in Google Shopping catalog construction using text features (using feed-forward RNN), image embeddings (used CNN trained with Triple Loss) and categorical features
    - Experimenting with Google Brain sentence embeddings to increase accuracy
    - Achieved F1 score of 0.88. Will be used in the construction/reconstruction of 2 billion catalogs daily
  - **Improving Description Quality in Display Data**: Replaced the manual curation process for two categories of quality issues with an algorithmic solution that both removed low quality descriptions and boosted descriptions that met transactability requirements
    - Launched improvements in production leading to improvements in transactability and quality of catalog descriptions for ∼100 million impressions/day
- **Cline Center for Advanced Social Research, UIUC** Urbana-Champaign, Illinois
  Research Assistant *Jan 2018 - May 2018*
  - **ResTeCo: Responsible Terrorism Coverage**: Used an ensemble of 3-gram contextual language model and a character level BiLSTM to correct OCR spelling errors (82% accuracy) on NY Times articles (Tensorflow, Keras)
    - Implemented LDA to analyze topics and LDA2Vec to extract topic features for classifying types of terrorist activities being reported in the data. Experimented with Contextual LSTM for topic prediction of news articles
- **University of Illinois at Urbana-Champaign** Urbana-Champaign, Illinois
  Research Assistant *Sep 2017 - Jan 2018*
  - **Stylometry in Research Papers**: Built a Mathematical model to identify style of writing research papers using statistical NLP techniques (in python, MATLAB)
- **Microsoft R&D India Pvt Ltd** Hyderabad, India
  Software Engineer *Jun 2015 - Jul 2017*
  - **OLS Simplification**: Migrated on-premise transaction system to Microsoft cloud infrastructure

Individual Projects

- **Curb Domestic Violence**: Ideated and used ensemble of LSTM on text (extracted from speech signals) and linear regression on change in decibel level of voice to predict domestic violence with a recall score 0.47
- **Lets Envisage**: Used Linear Regression to assign programming tasks to person based on their skill level in a team
  - Used Regression Analysis to predict the time required to complete a programming task based on its complexity with an RMSE of 4.5 hours

Presentations, Publication and Patent

- **Presenter**: S. Ghosh, S. Debnath. (Jun 2016). *Curb Domestic Violence.* Paper presented at Microsoft Machine Learning, Analytics & Data Science Conference, Redmond, 2016
- **Patent**: S. Ghosh, S. Debnath. (Nov 2016). *EFFICIENT TASK PLANNING USING PAST PERFORMANCE (Publication number: US20180150786A1).*

Achievements and Social Service

- **IT Star Award, Microsoft R&D India**: 2016
- **Semifinalist at ACM International Collegiate Programming Contest Asia Region**: 2014
- **Volunteering**: Teacher at SOS Children's Village an NGO for underpriviledged orphan children