PRANAV GOEL

Phone: +91 882 696 0840.
Email: pranavgoel403@gmail.com
Website: https://pranav-goel.github.io/

EDUCATION

Indian Institute of Technology (BHU) Varanasi

- Bachelor of Technology in Computer Science and Engineering (Graduated in 2018)
- Major: Computer Science with Honors in Artificial Intelligence
- Cumulative Performance Index (eight semesters): 9.34/10.0

Apeejay School, Pitampura, New Delhi

- All India Senior Secondary Board Examination 2014, CBSE (Percentage: 96.8%)
- All India Secondary Board Examination 2012, CBSE (Cumulative Grade Point Average: 10.0/10.0)

TECHNICAL EXPERIENCE

(common theme – intersection of Machine Learning and NLP)

Research Internship, Articulab, Language Technologies Institute, Carnegie Mellon University (Pittsburgh, May 2017 to July 2017)

- Worked on Conversational Strategy Classification for a Socially Aware Robot Assistant (SARA) with Dr. Yoichi Matsuyama and Prof. Justine Cassell.
- Automated detection of several social behaviors via the visual features of the user and the verbal features derived from the text of their utterance.
- Deep Neural Network based approaches using LSTMs and CNNs (utilizing Keras and Tensorflow libraries), a unique architecture for multi-label classification by combining LSTM/CNN with linear regression, and impact of leveraging dialog history were all explored.
- Outperforms previous classifier and overcomes other limitations of previous approaches, interesting insights into the nature of social strategies also gained.
- Error analysis for further interpretation ongoing, along with methods like transfer learning to improve performance.
- Did a spinoff research project on "Detecting Indirectness in Conversation". Work **accepted** at IWSDS 2018.

Software Engineering Intern, Ephesoft India (Gurugram, New Delhi NCR, December 2016 to January 2017)

- Worked with the Data Science team on content analysis.
- Fine-grained classification of new articles with proper comparison of approaches involving TF-IDF, Topic Modelling (using MALLET) and use of word vectors and document vectors (combined in various ways from word vectors or through Gensim functionality in Python).
- Microsoft's NLP tools were also leveraged. High performance of >95% accuracy achieved.

Research Internship, Center For Indian Language Tech. (CFILT) lab, Indian Institute of Technology Bombay (May 2016 to July 2016)

- Worked on a novel idea of automated detection of sarcasm target with Dr. Aditya Joshi and Prof. Pushpak Bhattacharyya.
- An important next step to sarcasm detection, proper definition of a 'sarcasm target' was crafted and its need in NLP explored.
- A hybrid system comprising of both rule based and SVM-HMM based classifiers was designed which improved upon two baselines (a naïve objective word-based approach and one inspired from sentiment target identification), and also outperformed the two individual classifiers..
- First system of its kind is tested on two datasets – tweets, and book snippets. Annotated by the project team (including me).
- The dataset has been released. Inter-annotator study to establish reliability led by me. Work has been **accepted** at LREC 2018.

INTERESTS AND SKILLS

- Interests: Intersection of Machine Learning and NLP, Neural Networks, Computational Linguistics, Distributional Semantics
- Programming Languages: Python, C++, C, R, MATLAB, Octave, SQL
- Tools and Technologies: Git, MySQL, Django, HTML, CSS, Latex, Gensim, Tensorflow, Keras, NLTK, Pandas, Numpy, SVM Perf, Mallet

PUBLICATIONS

- (Accepted, Camera-Ready to be submitted on June 9, 2018) - **Pranav Goel**, Devang Kulshreshtha, and Anil Kumar Singh. "How emotional are you? Neural Architectures for Emotion Intensity Prediction in Microblogs". COLING 2018. (2018)
- **Goel, Pranav**, Yoichi Matsuyama, Michael Madaio, and Justine Cassell. "I think it might help if we multiply, and not add": Detecting Indirectness in Conversation. In: Proceedings of the International Workshop Series on Spoken Dialogue Systems Technology (IWSDS). 2018.
- Joshi, Aditya, **Pranav Goel**, Pushpak Bhattacharyya, and Mark J. Carman. "Sarcasm Target Identification: Dataset and An Introductory Approach." LREC 2018-11th edition of the Language Resources and Evaluation Conference. 2018.

- **Goel, Pranav**, et al. "Prayas at EmoInt 2017: An Ensemble of Deep Neural Architectures for Emotion Intensity Prediction in Tweets." *EMNLP 2017*(2017): 58.
- **Goel, Pranav**, and Anil Kumar Singh. "IIT (BHU): System Description for LSDSem'17 Shared Task." *LSDSem 2017* (2017): 81.

SELECTED PROJECTS

<Research>

Predicting the correct ending of a story – system for the Story Cloze shared task at LSDSem workshop@EACL 2017 (paper published)

- Worked with Dr. Anil Kumar Singh on building a system which, given the context, could choose the correct ending to the story from two given options.
- Semantic similarity based approaches like using Word2Vec, Doc2Vec, Siamese LSTM, etc., their ensemble and sentiment similarity based approaches tried.
- Beat the baseline by 2%. A crucial negative result was also inferred. System description paper was accepted, and can be found in the workshop proceedings.
- Presented a poster in the LSDSem workshop to discuss the work and attended EACL 2017.

An Ensemble of Deep Neural Architectures for Emotion Intensity Detection in Tweets – EmoInt shared task at WASSA@EMNLP 2017

(paper published – Shared Task Winner at EMNLP 2017)

- Given a tweet and the emotion (out of anger, sadness, fear, joy) expressed in it the system predicts its intensity (real valued score from 0 to 1).
- Using pre-trained twitter word2vec embedding and affective lexicon based features, 3 approaches were tried – based on Feed-forward neural nets, Sequential Modelling using CNNs and LSTMs and neural multi-task learning. Predictions were combined.
- System ranked 1st out of 22 systems, setting the new state of art, beating task baseline by 10-15% Pearson correlation score across all emotions.
- System Description paper accepted, to appear in WASSA 2017 proceedings. Invited for oral presentation at WASSA@EMNLP 2017. Slides
- An extended version of this work – **accepted at COLING 2018**, which includes a *novel* neural architecture achieving even better results, enhanced deep multi-task learning approaches, the exploration of cognitive implications of correlations between pairs of emotions and error analysis has been submitted to TACL (Nov. 1, 2017).

Linguistic Regularities in Vector Space for Sentences

- Worked with Dr. Anil Kumar Singh on exploring the relationships exhibited by sentence vectors or embeddings, and if they show the same properties as word vectors (for example – capturing relationships like 'man' – 'woman' + 'king' ~ 'queen').
- Used sentence representations learnt from the large Stanford Natural Language Inference dataset.
- Discovered some interesting relationsips across multiple classes of a multi-class classification tasks (like between different emotions, or positive and negative sentiment classes) and investigated if these relationships can be captured by sentence embeddings.

<Academic>

Shallow Discourse Parsing

- Worked with Dr. Anil Kumar Singh on exploring the task of Discourse Parsing. Focused on the shallow version – extracting relations exhibited by connectives alone, which seems to have a remarkable scope for improvement in terms of performance.
- Implemented the current state of art pipeline system.
- Appending various submodules with semantic features and improving the pipeline structure was explored. Report

Automated Sarcasm Detection in Tweets

- Worked with Dr. Anil Kumar Singh on automated sarcasm detection on twitter tweets using hashtag based supervision.
- Successfully implemented the state of art at that time exploiting the linguistic theory of incongruity for sarcasm. LibSVM (RBF kernel) used for classification.
- Ablation tests to study impact of each category of feature (lexical, pragmatic, incongruity-based) were carried out. Slides

*The projects above are all related to my area of research interest – Machine Learning applied to NLP tasks. Other projects done by me during the course of my bachelors I. I.
:lude numerical algorithm implementations in Python and an interactive web app around the same, a blood bank database management system along with a web portal for the
stomer, a SQL query parser written in C++ and more.

TEACHING EXPERIENCE

- **Teaching Assistant** for the course 'Computer Programming in C' from January 2017 to May 2017.
- **Teaching Assistant** for the course 'Introduction to Artificial Intelligence' from January 2018 to May 2018.
- **Undergraduate Project Mentor** for exploratory projects, guiding sophomores on Natural Language Processing tasks with Prof. Anil Kumar Singh.

POSITIONS OF RESPONSIBILITY

- **Co-Convener** for Codefest 2017 – annual global programming festival conducted by Department of Computer Science and Engineering, IIT Varanasi. (September 2017)
- **Co-Coordinator** for the online world-wide ML competition MLWARE conducted within the annual techno-management festival of IIT Varanasi – Technex (February 2017)
- Member of the Club of Programmers, IIT (BHU) Varanasi. Responsible for all Machine Learning related initiatives.

ACHIEVEMENTS/EXTRA-CIRRICULAR

- Received INR 50,000 from institute as Travel Grant to attend EACL 2017.
- Selected for the DAAD-WISE 2017 summer research internship program.
- Secured top 50 position in world Machine Learning Codesprint 2016 - an online competition conducted at Hackerrank.com (handle – Pranav_Goel123).
- Started a Machine Learning Reading Group, focusing on applications in NLP and CV, for all undergrads interested in research in this field across the institute.
- Lead institute wide workshops for teaching the basics of Machine Learning and NLP to freshmen and sophomores.
- Participated and reached second round (top 160 teams) in a national level robotics events (theoretical designs and experiments on given robot) conducted by e-yantra in association with IIT Bombay (2016).
- Cleared IIT Advanced examination (2014) in first attempt with All India Rank 1051 (top 1%).
- Cleared Joint Entrance Examination 2014 for engineering conducted by CBSE with All India Rank of 360 (top 0.05%).