in anubhav-jangra

Anubhav Jangra

anubhav0603@gmail.com
+91-9812516727

Education

Bihar, India	IIT Patna	Jul 2017 – Jul 2021

- Bachelor of Technology in Computer Science and Engineering, GPA: 8.78/10, Major GPA: 9.11/10 (CS+MA)

Work Experience

Reserach Intern	GREYC Lab, ENSI-CAEN	Jun 2020 – Aug 2020

- Worked on patch-based identification of lexical semantic relation using PageRank-based attention mechanism.
- Adopted a novel multi-view framework to mimic the human perception using pattern-based, vision-based and cognition-based distributional features.

Research Intern	TCS Innovation Lab	Dec 2019 – Jan 2020

- Worked on a novel approach to tackle the insufficiency of data in time-series signals with Generative modeling.
- Achieved a 76.84% accuracy (72% without GAN) on Arrowhead dataset using proposed model (RESNET+GANs).

Research Intern	Kyoto University	May 2019 – Jun 2019

- Worked on a novel problem statement in the field of multi-modal summarization. Developed a new dataset.
- Formulated and implemented a novel joint-ILP framework that achieved 0.074 ROUGE-2 score for text summary, 59.9% precision and 38.3% recall for image summary, and 44% accuracy for video summary.

Research Intern	CFILT Lab, IIT Bombay	Dec 2018 – Jan 2019

- Learnt and implemented various Unsupervised NMT Deep Learning models for distant language pairs.
- Developed an unsupervised two-way translation system using WMT-2014 English-Hindi monolingual corpus.
- Exploited various techniques like random swapping of input words, and initializing the model using auto-encoder inspired contextual learning using attention based encoder-decoder models.

UG Research Scholar	AI-NLP-ML Lab, IIT Patna	Jul 2018 – Present

- Worked extensively in the area of summarization, varying from extractive and abstractive text summarization to multi-modal summarization with multi-modal outputs. Other explored areas include complaint mining and multi-label classification.
- Supervised four students (three interns and one junior) as a member of research lab.

Research Publications

- Anubhav Jangra, Adam Jatowt, Sriparna Saha, Mohammed Hasanuzzaman, "A Survey on Multi-modal Summarization". **ACM Computing Surveys 2021** (Impact Factor: **7.990**) (status: under review)
- Apoorva Singh, Sriparna Saha, Mohammed Hasanuzzaman, Anubhav Jangra, "Identifying Complaints based on Semi-Supervised Mincuts". **Elsevier's Expert Systems with Applications, 2021** (Impact Factor: **5.452**) (status: under review)
- Anubhav Jangra*, Raghav Jain*, Vaibhav Mavi*, Sriparna Saha, Pushpak Bhattacharyya, "Semantic Extractor-Paraphraser based Abstractive Summarization", **ICON 2020** (accepted)
- Anubhav Jangra, Sriparna Saha, Adam Jatowt, Mohammed Hasanuzzaman, "Multi-Modal Summary Generation using Multi-objective Optimization", **SIGIR 2020** (category **A*** conference)
  (DOI: `https://doi.org/10.1145/3397271.3401232`)
- Anubhav Jangra, Adam Jatowt, Mohammed Hasanuzzman, Sriparna Saha, "Text-Image-Video Summary Generation using Joint Integer Linear Programming", **ECIR 2020** (category **A** conference)
  (DOI: `https://doi.org/10.1007/978-3-030-45442-5_24`)
- Naveen Saini, Sriparna Saha, Anubhav Jangra, Pushpak Bhattacharyya, "Extractive Single Document Summarization using Multiobjective Optimization: Exploring Self-organized Differential Evolution, Grey Wolf

Optimizer and Water Cycle Algorithm", **Elsevier's Knowledge Based Systems, 2018** (Impact Factor: **5.921**) (DOI: `https://doi.org/10.1016/j.knosys.2018.10.021`)

Technical Experience

Projects

- **Multi-modal summarization (MMS)** (Jan 2019 - present) Have extensively explored this area for the past two years, and have extended this project as my undergraduate thesis.
  - First one ever to propose and solve the problem of text, image, and video summary generation.
  - Developed and implemented various systems using optimization techniques like integer linear programming, differential evolution, grey wolf optimizer etc. utilizing diverse objectives.
  - Formally defined the complementary and supplementary enhanced multi-modal summaries for the first time, and achieved a new state-of-the-art for the task of unsupervised MMS, surpassing the predecessor by almost twice as better ROUGE-2 score.
  - Submitted one of its kind literature survey at ACM Computing Surveys 2021.
- **Abstractive Text Summarization using reinforced learning** (May 2020 - present)
  - Proposed an 'extractor-abstractor' framework to outperform its predecessors by a margin of **0.5** ROUGE-1, **0.4** ROUGE-2, **1** METEOR, and **0.9** WMS (Word Mover Similarity) scores.
  - A knowledge discovery that the standard sequence-to-sequence networks like PGN model implicitly paraphrases was brought to light through this project.
  - Currently working on an adversarial model that is able combine information from multiple sentences into a latent space; overcoming a major shortcoming of sequence-to-sequence networks.
- **Complaint Mining** (Jun 2020 - Jul 2020)
  - An NLP task of binary classification of a review as a complaint using semi-supervised graph-based approach.
  - The proposed model surpasses the existing state-of-the-art by just utilizing only 50% of training data.
- **TensorFlow at the Edge for Wireless Drones** (Jan 2019 - Apr 2020)
  - Integrated edge computation by incorporating pixel-based motion detection on Raspberry Pi (mimicking the edge device) and object detection using in the intermediate nodes to detect human movement to reduce overloading on servers.
  - Received an AA (highest possible grade) in the CS299 course (Innovation Design Laboratory).
  - Project link: `https://github.com/dsciitpatna/tensorflow-at-the-edge`

Reviewer

- Reviewer at *ACM Computing Surveys*.
- Reviewer at *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP)*.
- Secondary reviewer at *European Association of Computational Linguistics (EACL) 2021*.

Additional Experience and Awards

- Google Research AI summer school (2020): Got selected to participate in the natural language understanding track of the summer school, limited to only 50 students amongst thousands of applicants.
- Attended a short term course and workshop on "Pragmatic Optimization for Practical Problem Solving" conducted by Michigan State university, IIT Roorkee and IISc Bangalore, limited to only 40 students.
- Department Lead ML, Google DSC IIT Patna (2019-2020): Supervised a few projects and gave lectures on Machine Learning theories and its applications.
- PyData Patna Conference (Dec 2020): Invited to give a talk on automatic text summarization.
- Ranked in National Top 0.2% (amongst 1,400,000 candidates) in JEE Mains 2017 and Top 1.5% (amongst 2,00,000 candidates) in IIT-JEE Advanced 2017.

Languages and Technologies

- Python, MATLAB, C/C++, MySql, PHP, Bash, HTML, CSS, JavaScript
- TensorFlow, Keras, PyTorch, Git, D3.js, Node.js, LaTeX, MongoDB