Syed Arbaaz Qureshi (he/him/his)

arbaaz.qureshi29@gmail.com ☎ +1 (413) 695-7408 linkedin.com/in/arbaazqureshi/ Syed Arbaaz Qureshi

Education

University of Massachusetts Amherst *Expected graduation: May 2023*
Master of Science in Computer Science. **GPA: 4.0/4.0**
Relevant courses: Machine Learning, Advanced Natural Language Processing, Probabilistic Graphical Models

Indian Institute of Technology Patna *August 2015 - May 2019*
Bachelor of Technology in Computer Science and Engineering. **GPA: 8.2/10**
Relevant courses: Machine Learning, Deep Learning, Natural Language Processing, Algorithms and Data Structures

Skills

Programming languages: Python, Java, C#, C
Software/Toolkits: TensorFlow, Keras, PyTorch, NumPy, SciPy, Pandas, Scikit-learn, MatPlotLib, Plotly, MySQL

Work Experience

Microsoft Research, India. Research fellow of team Sankie *August 2019 - August 2021*
• Advisors: Sonu Mehta, Dr. Rahul Kumar & Dr. Ranjita Bhagwan
• Projects:
1) ML for DevOps tasks involving code-edits

- Tasked with generating commit messages, and automatically classifying edit, based on the content of the commit.
- Scraped commit-commit message pairs from more than 1000 GitHub repositories. Built ACMG (Automatic Commit Message Generator), based on an existing work code2vec, and trained it on the scraped dataset.
- Built edit2vec (an extension of code2vec) for automatic code edit classification. Trained and analyzed it on the ManySStuBs4J dataset, and on another dataset that we collected.
- Achieved an accuracy of over 99% on code edit classification. Discovered code2vec isn't generalizable to other downstream tasks.

2) REX analysis

- Tasked with finding features in a repository to estimate the performance of REX, Microsoft Research India's pull request-level file recommendation system which mines file association rules from the commits of software repositories.
- Discovered repository features illustrated a Pearson correlation coefficient of 0.3 with REX precision.
- Made data-informed modifications REX, thereby improving its aggregate precision from 61% to 68%.

GREYC lab, University of Caen Normandy. Research intern of HULTECH team *May 2019 - July 2019*
• Advisor: Prof. Gaël Dias
• Project: **Analysis of the effects of gender on the estimation of depression severity [IJCNN 2021]**

- Developed various multitask learning model architectures to concurrently predict depression score in males and females, and trained them on the DAIC-WOZ dataset.
- Identified that a) gender information substantially improves the performance of depression severity estimation, and b) concurrently learning to predict depression level and the gender of the participant improves the performance of depression severity estimation.

AI-NLP-ML lab, IIT Patna. Undergraduate Research Assistant *August 2018 - May 2019*
• Advisors: Prof. Sriparna Saha, Prof. Gaël Dias & Prof. Mohammed Hasanuzzaman
• Projects:
1) Depression severity estimation using multimodal deep learning

- Designed and trained encoding networks for acoustic, text and visual modalities using the DAIC-WOZ dataset.
- Designed attention-based deep neural network for the fusion of modality representations obtained from the corresponding encoding networks.
- The proposed model has outperformed the state of the art by 7.17% on RMSE and 8.08% on MAE.

2) Multitask representation learning for multimodal estimation of depression severity [IEEE IS 2019]

- Developed different multitask learning model architectures to learn representations of individual modalities, by simultaneously predicting depression severity score and class.
- Outperformed the state of the art by 4.93% on RMSE and 1.50% on MAE. Set new baseline for depression classification, 66.66% accuracy and 0.53 F1-score.

3) Multitask learning to concurrently estimate emotion intensity and depression severity [IEEE CIM 2020]

- Designed and trained various multitask learning model architectures (fully shared, shared private and adversarial shared private) to concurrently predict depression score and emotion intensity using text data.
- Showed that substantial performance improvements in predicting the depression severity can be achieved by using emotion-aware models.

Selected projects

Research Paper Tagger (RPT), with Prof. Mohit Iyyer, at UMass Amherst.

Objective: Automatically tagging the research track of an NLP paper, given the title, abstract and the authors.

- Built a dataset of 1744 research paper-research track pairs from ACL 2021, 2020 and 2019. Fine-tuned a BERT-based classifier on the collected dataset, on various combinations of title, abstract and the authors of the research papers.
- Achieved a top-1 accuracy of 70% and a top-3 accuracy of 83% on the test split of the collected dataset. Working on publishing this work in a conference/workshop.

Cycle Location and Anti-Theft System (CLATS), with Prof. Jimson Mathew, at IIT Patna.

Objective: Tracking and preventing the theft of cycles inside gated regions.

- Built a working prototype of a lightweight system that detects the current location of the users cycle and prevents its usage without the consent of the owner.
- This project was awarded the most innovative project award by the faculty members of the CSE department.

Publications

1. **Gender-Aware Estimation of Depression Severity Level in a Multimodal Setting**
   Syed Arbaaz Qureshi, Gaël Dias, Mohammed Hasanuzzaman and Sriparna Saha.
   International Joint Conference on Neural Networks, 2021
2. **Improving Depression Level Estimation by Concurrently Learning Emotion Intensity**
   Syed Arbaaz Qureshi, Gaël Dias, Mohammed Hasanuzzaman and Sriparna Saha.
   IEEE Computational Intelligence Magazine, Volume 15, Issue 3, 2020
3. **Multitask Representation Learning for Multimodal Estimation of Depression Level**
   Syed Arbaaz Qureshi, Sriparna Saha, Mohammed Hasanuzzaman and Gaël Dias.
   IEEE Intelligent Systems, Volume 34, Issue 5, 2019
4. **Prédiction automatique des scores aux questionnaires PHQ-8 par intelligence artificielle**
   Gaël Dias, **Syed Arbaaz Qureshi**, Sriparna Saha and Mohammed Hasanuzzaman.
   French Journal of Psychiatry, Volume 1, Supplement 2, 2019

arXiv preprints

1. **Assessing the Effectiveness of Syntactic Structure to Learn Code Edit Representations**
   Syed Arbaaz Qureshi, Sonu Mehta, Ranjita Bhagwan, Rahul Kumar.
2. **The Verbal and Non Verbal Signals of Depression - Combining Acoustics, Text and Visuals for Estimating Depression Level**
   Syed Arbaaz Qureshi, Mohammed Hasanuzzaman, Sriparna Saha and Gaël Dias.