

Masum Hasan

PhD Student, Computer Science,
University of Rochester, NY
<https://masumhasan.net>

Research Interest

ML \cap Game Engines, Conversational AI, Synthetic Data

On-going Projects

- ✦ **SAPIEN: Synthetic Anthropomorphic Personal Interaction Engine**
 - SAPIEN is a platform for creating virtual humans with unique personalities and having human-like conversations with them.
 - Feedback to improve communication skills using NLP and audio processing
 - Made with Unreal Engine 5 Metahuman, GPT-3.5
 - Demo: [clinical communication training](#)
 - Demo: [language learning peer](#)
- ✦ **Hi5: Synthetic Data for Hand Tracking**
 - Synthetic data platform in Unity game engine
 - 550k precise hand pose estimation data for clinical hand tracking
 - Dataset demo [video](#)
 - Initial result [video](#)

Publications

- [CoDesc: A Large Code-Description Parallel Dataset](#). Masum Hasan, Tanveer Muttaqeen, Ishtiaq Niloy, Kazi Mehrab, Tahmid Hasan, Mahim Pantho, Wasi Ahmad, Rifat Shahriyar, Anindya Iqbal
Venue: **ACL Findings 2021**
- [Using a Balanced Scorecard to Identify Opportunities to Improve Code Review Effectiveness: An Industrial Experience Report](#). Masum Hasan, Anindya Iqbal, Amiangshu Bosu, Mohammad Rafid Ul Islam, A.J.M. Imtiajur Rahman.
Venue: **EMSE** (Journal), *Impact Factor: 8.41*
- [Review4Repair: Code Review Aided Automatic Program Repairing](#). Faria Huq, Masum Hasan, Mahim Pantho, Sazan Mahbub, Anindya Iqbal, Toufique Ahmed.
Venue: **IST** (Journal), *Impact Factor: 6.93*
- [Hitting your MARQ: Multimodal ARGument Quality Assessment](#). Md Kamrul Hasan, James Spann, Masum Hasan Md. Saiful Islam, Kurtis Haut, Rada Mihalcea and Ehsan Hoque.
Venue: **EMNLP 2021**
- [Not Low-Resource Anymore: Aligner Ensembling, Batch Filtering, and New Datasets for Bengali-English Machine Translation](#). Tahmid Hasan, Abhik Bhattacharjee, Kazi Samin, Masum Hasan, Madhusudan Basak, M. Sohel Rahman, Rifat Shahriyar.
Venue: **EMNLP 2020**
- [Recognition of Bengali Handwritten Digits Using Convolutional Neural Network Architectures](#). Md Mahmudul Hasan, Md Rafid Ul Islam, Md Tareq Mahmood.
Venue: **ICBSLP 2018**, ****Best Student paper award****
- [Early Detection of Earthquake Using Satellite Based Quantum Computing](#). Akhter Al Amin, Mahmudul Hasan, Kazi Sinthia Kabir, Tanzila Choudhury, ABM Alim Al Islam.
Venue: **ICCSNT, 2015**

On Archive!

- [Text2App: A Framework for Creating Android Apps from Text Descriptions](#). Masum Hasan, Kazi Sajeed Mehrab, Wasi Uddin Ahmad, Rifat Shahriyar.
Project page: <https://text2app.github.io/>

- [Auto-Gait: Automatic Ataxia Risk Assessment with Computer Vision on Gait Task Videos.](#) Wasifur Rahman, **Masum Hasan**, Md Saiful Islam, Titilayo Olubajo, Jeet Thaker, Abdelrahman Abdelkader, Phillip Yang, Tetsuo Ashizawa, Ehsan Hoque.
Under review: **UbiComp 2022**
- [BERT2Code: Can Pretrained Language Models be Leveraged for Code Search?](#) Abdullah Al Ishtiaq, **Masum Hasan**, Md. Mahim Anjum Haque, Kazi Sajeed Mehrab, Tanveer Muttaqueen, Tahmid Hasan, Anindya Iqbal, Rifat Shahriyar

Professional Experience

- Graduate Research Assistant, ROC-HCI, University of Rochester, NY** *Aug 2021 to Present*
- ⊕ Supervised by Dr. Ehsan Hoque
- Research Assistant (RA), Applied Machine Learning Lab, BUET** *Nov 2018 to Jul 2021*
- ⊕ Supervised by, [Dr. Rifat Shahriyar](#), [Dr. Anindya Iqbal](#)
 - ⊕ ML for source code: Source Code Summarization/Synthesis/Search/Repair/Compression
 - ⊕ NLP for Bengali: GPT-2, Bengali-English NMT
- Cofounder and CEO, Bizzy Ltd.** *Dec 2016 to Jul 2017*
- ⊕ [BizzyBd](#) started as an easy-to-use website building tool for non-programmer professionals.

Education

PhD Student, Computer Science (CS)

University of Rochester (UR) ~ Started Aug 2021

BSc. In Computer Science and Engineering (CSE)

Bangladesh University of Engineering and Technology (BUET) ~ Oct 2018

- ⊕ Thesis: Mythbusting Gender Stereotypes Using Natural Language Processing

Online Courses

- ⊕ [Machine Learning](#), Coursera.
- ⊕ [Deep Learning Specialization](#), Deeplearning.ai
- ⊕ [Practical Deep Learning for Coders](#), Fast.ai
- ⊕ [Book] [Neural Network and Deep Learning](#), Michael A. Nielsen
- ⊕ [Partial] [CS224n: Natural Language Processing with Deep Learning](#), Stanford University
- ⊕ [Essence of Linear Algebra](#), 3blue1brown

Blogs and Articles

Machine Learning:

- [Absolute Beginner's Guide to Machine Learning and Deep Learning](#)
- [Setup and Run fast.ai in Amazon AWS](#)

Graduate Admission:

- [How I got 4 Ph.D. offers in the US with a CGPA 2.79 — and what you can learn from it](#)

Extracurricular Activities

- ⊕ Created an introductory [Machine Learning](#) course in my native language Bengali that dives into theory, practice, and intuition behind neural networks.
- ⊕ Leisure: reading, improv, taekwondo, meditation, onewheel, existential dread