

# AYUSH KUMAR SHAH

Ph.D. student in Computer Science

✉ as1211@rit.edu    📍 Rochester, New York    📞 (585) 471-9866  
🌐 ayushkumarshah    🔗 linkedin.com/in/ayush7/    🖥️ shahayush.com

**Research Interests:** Pattern recognition, computer vision, detection and recognition of graphical structures, multi-modal deep learning, natural language processing, visual scene parsing

## WORK EXPERIENCE

---

### Amazon - Alexa AI

*Applied Scientist Intern*

Sunnyvale, California

*May 2022 – Aug 2022*

- Working on Alexa Perceptual Technologies - AI team

### Fusemachines

*Machine Learning Engineer*

Kathmandu, Nepal

*June 2019 – Aug 2020*

- Developed a product classifier using chemical attributes to optimize business decisions for products that go unsold using boosting algorithms including Gradient Boosting, Random Forests, XGBoost, LightGBM.
- Built an intelligent character recognition system using CNN and RNN to predict handwritten texts (both English and Nepali) in manually-filled form fields with an accuracy of 95%.
- Analysed data provided by a subscription-based e-commerce client for building a recommendation system, which led to an increase in revenue through cross-selling by 6% (large as we were serving 600k users).
- Worked on preparing course materials for Fusemachines AI Education Programs - “Foundations in AI: Mathematics for AI” , “Micro Degree™ in Artificial Intelligence: Machine Learning, Computer Vision”

### Samriddhi College

*Computer Science Instructor*

Kathmandu, Nepal

*Jan 2020 – June 2020*

- Designed and implemented daily lesson plans and coding sessions for the course “Foundations in AI: Computer Science and Mathematics” to undergraduate BSc.CSIT students. The course topics include: Introduction to AI, Fundamentals of CS, Python Programming, Data Structure, Database Management System.

## EDUCATION

---

### PhD in Computing and Information Sciences

Rochester Institute of Technology (RIT)

CGPA: 3.92/4

**Advisor:** Dr. Richard Zanibbi

**Research Group:** Document and Pattern Recognition Lab (DPRL)

*Aug 2020 – Present*

*Rochester, NY, USA*

**Area of focus:** extraction and visual parsing of graphical structures and notations, focusing on mathematical formulas and chemical diagrams in documents.

**Relevant Courses:** Pattern Recognition, Computer Vision, Mathematics for Deep Learning, Natural Language Processing, Software Engineering

### Bachelors in Computer Engineering

Kathmandu University

CGPA: 3.96/4

**Relevant Courses:** Artificial Intelligence, Data Structures and Algorithms, Algorithm and Complexity, Software Engineering, Probability and Statistics, Machine Learning, Speech and Language Processing, C, C++

*Aug 2015 – Oct 2019*

*Kavre, Nepal*

## PUBLICATION

---

- A. K. Shah**, A. Dey, and R. Zanibbi, “A Math Formula Extraction and Evaluation Framework for PDF Documents,” in Document Analysis and Recognition – ICDAR 2021, Cham, 2021, pp. 19–34. doi: 10.1007/978-3-030-86331-9\_2

## RESEARCH EXPERIENCE

---

### Document and Pattern Recognition Lab (DPRL), RIT

Graduate Research Assistant

Rochester, New York

Aug 2020 – Present

- Worked on the MathSeer project, a system to make finding mathematical information easier by creating innovative search engines, interfaces, and algorithms for extracting and recognizing math
- Built a new open-source math formula extraction pipeline for PDF files
- Adopted distributed parallelization methods with multiple GPUs and implemented custom dataloader with dynamic batch size to fully utilize the GPU, which increased the speed of the math formula parser by 6 times
- Built new tools for visualization and evaluation of parsing results and errors
- Worked on a PDF symbol extractor that identifies precise bounding box locations in born-digital PDF documents
- Developed a simple and effective algorithm to perform detection of math expressions using visual features alone
- Wrote an API for recognizing handwritten and typeset formulas and output the corresponding  $\LaTeX$  and MathML
- Currently working on improving the accuracy of the math formula parser by experimenting better visual features and attention mechanisms
- Currently working on adopting the parser to work with more complex graphical structures like chemical diagrams

## TECHNICAL SKILLS

---

<b>Programming Languages</b>	Python, R, Matlab, C, C++, JAVA
<b>Python Packages</b>	Pytorch, Tensorflow, Scikit-Learn, OpenCV, Nltk, Pandas, Numpy, Matplotlib, Fastapi, BeautifulSoup, Regex, NetworkX, Jupyter
<b>Database</b>	MySQL, MongoDB
<b>Miscellaneous</b>	Git, Github, Bash, $\LaTeX$ , Jira, Linux, Arduino, Raspberry-pi

## HONORS AND AWARDS

---

- RIT Ph.D. Merit Scholarship/Assistantship.** Financial Support for Ph.D. at RIT. 2020 – Present
- Kathmandu University Merit-based scholarship (4x).** \$440 worth scholarship awarded for securing the highest GPA in the Computer Engineering cohort (4/7 semesters). 2015 – 2019
- Fusemachines Artificial Intelligence Scholarship Program.** Selected among thousands of candidates nationwide for fuse.ai Artificial Intelligence Scholarship Online Course. Nov 2018
- American Society of Nepalese Engineers Merit Award.** A merit worth \$200, rewarded to the entrance topper of each university in Nepal, seeking admission for undergraduate degrees. May 2016
- 46<sup>th</sup> International Physics Olympiad (IPhO) Contestant.** One of the largest olympiads for high school Physics enthusiasts with 5 contestants, each from 100 participating countries. June 2015

## PROJECTS

---

- Nepali Plagiarism Detector** 2019  
An application that detects plagiarised Devanagari text files using a self-built rule-based stemming algorithm and Cosine similarity.
- Guitar chord recognizer** 2019  
An application that predicts the chords when the Mel spectrograms of guitar sound are fed into a CNN.
- AI Plays GTA 5: Simulating self-driving vehicles** 2019  
A bike-riding agent in a virtual environment (GTA5), built using CNN, used for simulating self-driving vehicles.
- Sarangi: Nepali lyrics emotions extraction** 2018  
A framework that categorizes songs written in the Devanagari script into four emotions using Naive Bayes.
- AutoCar** 2018  
A self-driving car that can detect lanes, stop sign, traffic light and avoid a collision, built using Canny edge detection, Hough transform, Haar cascade classifier, and Arduino programming.
- MathMate – advanced mathematical calculator** 2018  
An android app that solves different types of mathematical equations, numerical computations, and calculus problems showing involved steps.