Akshat Choube

akshatchoube3011@gmail.com| akshat3011.github.io | linkedin.com/in/akshat-choube | 213-800-6697

RESEARCH INTERESTS

I am interested in using machine learning and human-computer interaction techniques to study and build "behavior-aware" intelligent systems in the healthcare and education space.

EDUCATION

University of Southern California, Los Angeles **GPA: 3.95/4.0** **August 2019 – May 2021**
Master of Science in Computer Science (Honors)

Indian Institute of Technology (IIT), Palakkad **GPA: 8.89/10** **August 2015 – April 2019**
Bachelor of Technology in Computer Science and Engineering

WORK EXPERIENCE

Software Development Engineer (ML) | Amazon Search **July 2021 – present**
Using ML and NLP to understand shopping intents in customer queries on the Amazon website
- Worked on different NLP methods to increase query parsing coverage for a Gazette-based Ontology Linker.
- Generated distribution parameters for price and shipping speed of products shown for search queries based on customer behavior.
- Worked on understanding media intent in queries using a multi-lingual multi-intent BERT-based model.
- Working on an experiment to enhance customers' autocomplete experience by showing related keywords for a typed prefix.

Graduate Research Assistant | Intelligent Human Perception Lab, USC **Jan 2021 – May 2021**
Worked on predicting mood using apple watch data
- Worked on mood prediction using heartbeat and exercise data from apple watch.
- Performed statistical tests and used deep learning models to understand the correlation between self-reported mood and Apple watch data.

Teaching Assistant | Viterbi School of Engineering, USC **Jan 2021 – May 2021**
TA for Applied NLP (CSCI-544)
- Conducted office hours to help students with NLP theory, projects, and report writing.
- Corrected students' test papers and held sessions to brainstorm ideas for their projects.

SDE Intern | Amazon Search **June 2020 – Aug 2020**
Spark-Based Data Processing Framework
- Designed and implemented a framework to convert multiple offline user behavior and clickstream datasets into a read-only database (RODB) which is deployed to the online search service.
- The framework provides custom data preprocessing options like filtering, transformation, aggregation, and merging.
- The RODB file generated allows fast access and reduces overall search latency.

Graduate Research Assistant | Social Media Analytics Lab, Keck School, USC **Sept 2019 – Jan 2021**
Using NLP for understanding public health behaviors
- Analyzed Twitter for conversations, marketing, and misinformation around tobacco products (like cigarettes, cigars, etc.) and cannabis products (like weed, hash, etc.).
- Analyzed the roles of social bots in spreading misinformation and promoting hazardous products.
- Conducted analyses on the predominant health effects of cannabis use on Twitter.

Summer Intern | Divum Labs **May 2018 – July 2018**
Using Blockchain for e-commerce
- Modified Merkle Patricia trie in Ethereum framework to allow storing e-commerce receipts in the blockchain.
- Represented ecommerce data from blockchain as Knowledge Graph and applied machine learning algorithms on it.
- Compared Graph databases like Neo4j and Dgraph for performances on read, write, and query search operations.

Junior Research Fellow | Tata Institute of Fundamental Research **May 2017 – Dec 2017**
Information exchange system involving bit packets

- Defined a “Quality of Service” metric considering transmission time and the fraction of a bit packet.
- Proved optimal order for transmission of packets and devised a greedy algorithm that performs at most twice worse than the intractable optimal solution.
- Published and presented this work at National Conference on Communications (NCC 2018) (link)

PUBLICATIONS

[1] **Akshat Choube**, Mohammad Soleymani, “Punchline Detection using Context-Aware Hierarchical Multimodal Fusion”, International Conference on Multimodal Interaction (ICMI), 2020.

[2] Rahul Vaze, **Akshat Choube**, Shreyas Chaudhari, Nitin Aggarwal, “Energy-Delay-Distortion Problem”, National Conference on Communications (NCC), 2018.

[3] Jon-Patrick Allem, Allison Dormanesh, Anuja Majmundar, Jennifer Unger, Mathhew Kirkpatrick, **Akshat Choube**, Aneesh Aithal, Emilio Ferrara, Tess Cruz, “Topics of Nicotine-Related Discussions on Twitter: Infoveillance Study”, Journal of Medical Internet Research (JMIR), 2021.

PROJECTS

Persona and Emotion-Aware Dialogue System
- Created a new dataset by augmenting the PERSONA-CHAT dataset with discrete emotions for utterances using an automatic emotion classifier network (Deepmoji).
- Finetuned GPT-2 on the new dataset to generate a response given the agent's persona, dialogue history, and desired emotion.
- Analyzed the interplay between the agent’s emotion and persona in different dialogue scenarios by varying intensity and label of emotions.

Multimodal Punchline Detection
- Proposed a context-aware hierarchical network for multimodal punchline detection in TED talks.
- Demonstrated the ability of hierarchical fusion to learn multimodal interactions for humor using an ablation study.
- Achieved state-of-the-art accuracy on UR-FUNNY Dataset and published the work at ICMI 2020 (link).

Analyzing Happy Moments
- Designed and implemented a Sentiment Analysis study to understand sources of happiness among people of different cultures, sex, and age based on HappyDB dataset containing 14K happy moments.
- Generated word clouds to visualize the data and classified happy moments into 7 categories using LSTM achieving 81% accuracy.
- Developed a valence scoring algorithm for sentences considering contextual valence shifter words (like not, barely, etc.)

Multiview Face Synthesis using Generative Adversarial Network (GAN)
- Studied, implemented, and compared existing Multiview face generation models like CR-GAN, DR-GAN, TP-GAN, and GANnotation.
- Implemented a distributed version of CR-GAN using PyTorch to train in a multiple CPU/GPU environment.

Intelligent Document Summarizer
- Created a Python application that summarizes a large text document into a few sentences that best describes it.
- Devised an algorithm that clusters sentences based on their tf-idf vectors and then includes the sentence with the maximum average cosine similarity score within the cluster in summary.

Expert Paper Recommendation System
- Designed a recommendation system that automatically assigns papers to experts for conference reviews.
- Developed an algorithm that learns an expert's domain of work from their previous papers and recommends papers in the same domain.

TECHNICAL SKILLS

- **Programming Languages:** Python, C, C++, Scala, Haskell, Java, HTML, CSS
- **Database:** MySQL, MongoDB, Dgraph, Neo4j
- **Frameworks/Tools:** PyTorch, NLTK, Spark, Keras, TensorFlow, Visual Studio, LaTeX, Git, Matlab, ArcGIS

EXTRA-CURRICULAR ACTIVITIES

- Managed corporate relations for Dept. of Computer Science, IIT Palakkad and secured 100% recruitments.
- Conducted practical sessions on Machine Learning for 70+ teachers as part of the Additional Skill Acquisition Programme.
- Taught 50+ underprivileged students as part of National Service Scheme (NSS) and motivated them to pursue science.