Sina Ehsani - sinaehsani@arizona.edu - (+1) 310 874 6216

github.com/sina-ehsani - linkedin.com/in/sina-ehsani - sina-ehsani.github.io

EDUCATION

University of Arizona

Ph.D. in Systems and Industrial Engineering, MultiModal Deep Learning	*2023*
M.S. in Computer Science, NLP	*2022*
M.S. in Systems and Industrial Engineering, ML	*2018*

TECHNICAL STRENGTHS

Computer Programming	Python, Java, Scala, SQL, Bash, R
Software & Tools	Docker, MlFlow, WandB, Git, Scikit-learn, NLTK, spaCy, AWS, AzureML
Deep Learning Frameworks	PyTorch, TensorFlow, HuggingFace, Keras, PyTorch Lightning

RELEVANT COURSES

Design and Analysis of Algorithms, Algorithms for Natural Language Processing, Neural Networks, Applications of Machine Learning, Principles of Machine Learning, Artificial Intelligence for Health and Medicine, Text Retrieval And Web Search, Computational Linguistics, Software Engineering, Advanced Operating Systems, Engineering Statistics, Fundamentals of Optimization, Stochastic Modeling, Applied Cyberinfrastruc, Enterprise Data Management

WORK EXPERIENCE

Data Scientist, American Airlines *Aug 2022 - Present*

· *Developing* a **deep neural network forecasting engine** that predicts flight-level traffic for the Yield Management team with improved accuracy of more than 80% compared to previous statistical and machine learning frameworks. Reducing prediction mean square errors by 50% through the use of multimodal neural network architecture and pre-training on historical data, including time series and multi-dimensional data.

· *Leading* the development of the forecasting engine and spearheading its deployment into the airline system for A/B testing and potential replacement of older methods. **Training and supervising** a team of four in the use and evaluation of the approach.

· Expanding the analysis from one hub to cover the entire market of around 7,000 flights, and planning to achieve accurate predictions for future 330 days. Demonstrating the model's superior performance in capturing trends and seasonality, as well as its ability to observe sudden changes in capacity effects on demand that previous models were unable to capture.

· Improving the model's performance through the addition of more data, such as training the model on routes with multiple flights per day compared to those with fewer flights.

Machine Learning Research Intern, Truveta *May 2022 - Aug 2022*

· Contributed to the company's mission of Saving Lives with Data through the development of data-driven NLP approaches.

· Pre-trained and fine-tuned transfer-based models on clinical text data, resulting in improved overall accuracy for upstream and downstream tasks and significantly increased downstream training and inference speeds (fourfold and twofold increase, respectively).

· Achieved state-of-the-art results in the ontology alignment task, with at least a 5% improvement in F1, Hit@1, and MRR compared to previous works. Co-authored a paper proposing a novel framework, Truveta Mapper (TM), which leverages a multi-task sequence-to-sequence transformer model for unsupervised ontology alignment. TM outperforms existing solutions in terms of runtime latency and alignment quality, offers log-linear complexity, and makes the ontology alignment task more efficient and straightforward without extensive post-processing involving mapping extension or repair.

Data Science Research Intern, RedShred *Mar 2021 - Aug 2021, Jan 2022 - May 2022*

· Implemented a novel plot-processing approach utilizing object detection models and text extraction techniques, resulting in a 69.12% success rate in correctly identifying data from plots, **approaching the state-of-the-art** in the field. Deployed on the RedShred API for automatic processing of PDF documents.

· Utilized HuggingFace framework to develop and train transformer-based methods for document understanding, including segmentation and object detection. These techniques demonstrated strong performance in the field of document analysis.

Graduate Research Assistant, University of Arizona *May 2018 - Dec 2022*

Deep-Shallow Network:

· Developed a new deep learning multimodal structure that outperformed baselines such as RNN, CNN, CNN-LSTM, and ConvLSTM models as well as statistical models with 5% better MSE loss in predicting NYC taxi-traffic, passenger flight, and transportation data.

· Demonstrated the model's ability to capture recent changes in trend and special events in data, leading to statistically significant results.

Multimedia Systems for Textual Question Answering:

- Implemented advanced neural networks including transformers (BERT, Roberta), CNNs (ResNet), and RNNs (LSTM) for multimodal systems to capture spatial and sequential relations of data.
- Improved accuracy on multiple benchmarks through the use of on-demand image augmentation from the web, resulting in statistically significant improvements in accuracy for textual question answering.
- Conducted an interpretability study demonstrating the use of visual grounding for better decision-making by the model.

Unsupervised Machine Learning for Traumatic Brain Injury:
- Applied subspace clustering techniques to a dataset of 882 TBI patients, resulting in the identification of three clusters of laboratory physiological data.
- Demonstrated the feasibility and applicability of subspace clustering methods for phenotype definition and validation in the clinical domain of TBI, providing a foundation for further research and validation.

Teaching Assistant:
- Assisted in teaching courses on Introduction to Engineering Probability and Statistics, Survey of Optimization Methods, and Statistical Quality Control, providing support to students and assisting with course materials.

PROJECTS

Deep Learning and Natural Language Processing Projects

- *Music Generator*: Deep neural networks were used to create a music creation tool that automatically generates, extends, and stylizes imported MIDI music compositions.
- *Mental State of Tweeter*: Recurrent neural networks (e.g. GRU, LSTM) were used to create a multi-label classification model that achieved top ten in the CodaLab competition.
- *Domain Adaptation in Reading Comprehension*: Applied Domain Adaptation, LSTM, BiRNN, GRU, FFN, and Attention units with PyTorch on the SQUAD Question Answering task.
- *Recognizing Textual Entailment on SNLI Corpus*: Used a straightforward but effective Naive Bayes machine learning approach on the Stanford Natural Language Inference (SNLI) challenge, producing results competitive with those of neural network models.

Database, Data Visualization, and Web Development Projects

- *Design and Implementation of a Web-Based Database Application*: Worked on the front-end and back-end of a database system that can assist a school class management with its day-to-day operations and decision-making. The database system was built using Oracle SQL Server and web-based development tools.
- *Programming Distribution Maps for Butterfly Species*: Collaborated with 12 team members to collect and visualize the distribution and monthly migration maps for 700 North American Butterfly species using University of Arizona High-Performance Computing (HPC) servers, Docker, and Singularity Containers.

PUBLICATIONS

S. Ehsani, J. Liu (2023). OD-TQA: On-Demand Visual Augmentation for Textual Question Answering Task. Submitted to ACM SIGIR 2023.

M. Amir, M. Baruah, M. Eslamialishah, *S. Ehsani*, A. Bahramali, S. Naddaf-Sh, and S. Zarandioon (2023). Truveta Mapper: A Zero-shot Ontology Alignment Framework. Submitted to the IJCAI 2023.

S. Ehsani, C. K. Reddy , B. Foreman, J. Ratcliff, & V. Subbian (2021). Subspace Clustering of Physiological Data From Acute Traumatic Brain Injury Patients: Retrospective Analysis Based on the PROTECT III Trial. JMIR Biomedical Engineering, 6(1), e24698.

PROFESSIONAL AND PERSONAL DEVELOPMENTS

Manager of the international students' sports activities club (ICAT) at the University of Arizona.	2022
Manager of UA's Computational Medicine and Informatics for Neurological Health (COM-IN) Lab.	2018
Completed Six Sigma Green Belt training and Primavera Project Management certification.	2016
Second place in the state tournament for the college basketball championship, Isfahan, IR.	2014