Ali Naeimabadi

Natural Language Processing, Data Science, Databases, Deep Learning

linkedin.com/in/naeim-ali github.com/anaeim +1 (438) 924-0342 naeimaba@ualberta.ca personal webpage

Database Lab, Computing Science Department, University of Alberta, Edmonton, Canada, T6G 2R3

Education

2020-2023 **University of Alberta**, Edmonton, Canada GPA:3.84/4

M.Sc. in Computing Science

Courses : Modern Database management Systems, Introduction to NLP, Machine Learning, Deep Learning for NLP, Intro to Information Retrieval, Representation Learning at MILA, Introduction to Knowledge Graphs, Topics in Data Bases

Thesis : Product entity matching by leveraging pre-trained langauge models and tabular data ; Supervisor : *Prof. Davood Rafiei*

2011-2016 **Amirkabir University of Technology**, Tehran, Iran GPA:3.96/4

B.Sc. in Computer Engineering

Related courses : Artificial intelligence, Foundations of language and speech processing, Foundation of data mining, Data structure and algorithms, Software engineering, Data base, Engineering statistics and probability

Thesis : Microscope images processing to predict fracture strength of nanoclay/polyamide 12 nanocomposite using CNN ; Supervisor : *Dr. F. Hemmati*

Research interests

NLP	High-dimensional semantic similarity search, Question answering, large language models, entity matching, user feedback analysis, sentiment analysis, information retrieval
Database	relational databases, no-SQL databases, query expansion, query prediction
Data science	Data mining, database integration, entity matching, text mining, big data, database query
Artificial intelligence	Machine learning, deep learning, few-shot deep learning, attention models, transformers, BERT

Projects

E-commerce product matching model using Pre-trained Language Models (PLM) and tabular data : TATEM

A fast, highly informed and effective record-matching model is designed to find the matching pairs among online e-commerce platforms : Google-Amazon and Walmart-Amazon. For the first time, the model benefits from product titles and other complementary textual and tabular information, *supervised by Prof. Davood Rafiei*

entity matching | record matching | pre-trained language models | table serialization | tabular data encoding | high-dimensional similarity search | web tabular data extraction | BERT | RoBERTa | Universal Sentence Embedding | deep learning | PyTorch | Beautiful Soup

Introduced two product entity matching datasets using complementary product-specific tabular data : Amazon*-Google, Walmart*-Amazon*

For the first time Online product entity matching datasets include product-specific tabular data as complementary domain knowledge, *supervised by Prof. Davood Rafiei*

tabular data | entity matching | online product matching | Amazon-Google dataset | Amazon-Google-Tab dataset | Amazon-Google-Tex dataset | Beautiful Soup

Fast and robust domain adaptation for automated essay scoring

The data-efficient cross-domain model was designed to adapt knowledge from source domains to unseen target domain in low-resource scenarios using meta-learning

meta-learning | few-shot learning | 2-way 5-shot text classification | prototypical networks | induction networks | educational assessment | Automated essay scoring | Glove-word embedding | neural networks | hand-crafted linguistic features | PyTorch | text analysis tools | coh-metrix | SiNLP | TAALED | TAALES | TAACO | GAMET

Few-shot entity linking in knowledge graphs using limited resources

The model encoded the entities descriptions by BERT-based transformers and executed high-dimensional similarity search to find promising candidates for an unseen knowledge graph in zero-shot setting, *supervised by Dr. Denilson Barbosa*

knowledge graphs | BLINK | similarity search in high-dimensional space | FAISS indexer | wikipedia | wikidata | deep learning | transformers | BERT | PyTorch

Automated essay scoring model based on cohesion, coherence and text diversity linguistic features

A deep neural networks model was designed to predict the score of essays using deep hand-crafted features, *supervised by Dr. Denilson Barbosa*

educational assessment | automated essay scoring | deep learning | PyTorch | cohesion | coherence | text diversity | grammar error | mechanical error | SALAT | text analysis tools | coh-metrix | SiNLP | TAALED | TAALES | TAACO | GAMET

Database integration between two large-scale databases at ScotiaBank

Database integration between two large internal and external databases to enrich the client's information through a 2-stage string matching method : LSH and a Snorkel rule-based classifier, *supervised by Dr. Denilson Barbosa*

databases | database integration | pyspark | hadoop | sql | machine learning | locality sensitive hashing (LSH) | snorkel

Enhanced deep learning models for automated essay scoring

The project studied the effect of type of word vector representations on the accuracy of automated essay scoring systems, *supervised by Dr. Carrie Demmans Epp and , and Dr. Denilson Barbosa*

NLP | language model | automated essay scoring | LSTM | RNN | CNN | deep learning | TensorFlow | keras | BERT | Word2Vec | GloVe | pre-trained GloVe | pre-trained word embedding | nltk | gensim | python

Multi-class sentiment analysis using deep learning

The project studied sentiment analysis using deep learning techniques on Digikala website reviews to measure customer satisfaction, *supervised by Dr. Hossein Zeinali*

NLP | sentiment analysis | deep learning | machine learning | pytorch | CNN | RNN | SVM | Naive-Bayes | decision tree | nltk | python

Developing an AI framework for microscope image processing

The project performed microscope images processing to predict fracture strength of nanoclay/polyamide 12 nanocomposite using VGG16, *supervised by Dr. F. Hemmati*

image processing | semantic segmentation | deep learning | CNN | VGG16 | tensorflow | python

Developing a deep learning model to predict chemical reaction

A deep learning model was used to predict atomic energies of lignin-based bio-lubricants as a measure of thermal stability using CNN, *supervised by Prof. H. Garmabi*

reaction simulation | chemical engineering | process optimization | deep learning | k-bag | CNN | TensorFlow | python

Computer skills

Programming Languages	Python, C++, java, MATLAB, Scala
Operating Systems	Ubuntu, Windows, MacOS
Programing	PyTorch, TensorFlow, Keras, scikit-learn, SciPy, NumPy, Pandas, NTLK, GENSIM, Seaborn, Matplotlib PostgresSQL, MySQL, PySpark, Neo4j

Honors and awards

2022	Awarded Alberta Innovates Graduate Student Scholarship (26,000 CAD), Alberta Innovates
2020	Awarded full scholarship (54,000 CAD) for MSc. program, the University of Alberta
2019	Honored to win the Excellence in Scientific Achievement Award at Faroob Zaman Inc.
2019	Chosen as a member of national elite foundation of Iran
2019	Ranked 8th in national university entrance exam among Master applicants of Iran in computer engineering
2018	Received Scholarship from Iranian Nanotechnology Initiative Council
2016	Awarded full scholarship from Amirkabir University for MSc program at computer engineering department without taking entrance exam, granted to top 1% of the students
2016	Ranked 3rd among students of computer engineering department and graduated with honors, Amirkabir University of Technology in class of 2011
2011	Ranked among top o.1% in national university entrance exam among over 700,000 participants
2011	Ranked 3rd in Khayyam math Olympiad in Khorasan province
2010	Ranked 1st in Khayyam math Olympiad in Khorasan province

Working Experiences

Sep 2021 – Apr 2022 | **Research assistant at AML/ATF department, ScotiaBank, Edmonton, Canada**

Supervisor : Dr. Denilson Barbosa

- Database cleaning
- Database integration using a 2-stage string matching methodology
- Implement the solution into the pipeline

Jan 2021 – Apr 2021 | **Teacher assistant for search, knowledge, simulation, University of Alberta, Edmonton, Canada**
Supervisor : Dr. James Wright
- Helping the teacher with grading, answering the questions on the course forum and other administrative tasks

Sep 2020 – Dec 2020 | **Teacher assistant for games, puzzles and algorithms, University of Alberta, Edmonton, Canada**
Supervisor : Porf. Ryan Hayward
- Marking exams and quizzes, preparing variant of questions and designing new questions, answering student questions

Sep 2016 – June 2020 | **AI scientist at Faroob Zaman Inc., Mashhad, Iran**
Faroob Zaman, a start-up incubator active in AI, NLP and data science
- Designing AI framework for sentiment analysis on a website reviews to evaluate customer satisfaction
- Developing AI-assisted microscope image processing software for thin-layer lab, University of Tehran
- Designing a deep learning model to predict the yield of non-linear chemical reactions at Tehran refinery. It help perform a global search for new lubricant products

Sep 2015 – Dec 2015 | **Teacher Assistant, Engineering statistics and probability, Amirkabir University, Tehran, Iran**
- Grading assignment, Assist teachers with lesson preparation by getting materials ready

Sep 2014 – Dec 2014 | **Teacher Assistant, Advanced Programming, Amirkabir University, Amirkabir University, Tehran, Iran**
- Attending all training classes, providing support and guidance to students in weekly office hours

Language skills

- Persian — Native
- English — Professional

Affiliations

- University of Alberta
- Amii

Publications

Proceedinds :

- **A. Naeim abadi**, T. Nayeem, D. Rafiei, TATTOO : Product Entity Matching as a Topology Construction, WSDM 2024 [submitted].
- **A. Naeim abadi**, T. Nayeem, D. Rafiei, Product Entity Matching via Tabular Data, CIKM 2023.
- T. Firoozi, O. Bulut, C. Demmanse Epp, **A. Naeim abadi**, D. Barbosa, The Effect of Fine-tuned Word Embedding Techniques on the Accuracy of Automated Essay Scoring Systems Using Neural Networks, NCME 2022, April 21-24, 2022, San Diego, USA.
- **A. Naeim abadi**, An introduction to the radio communication systems, 1st pulse magazine on electrical engineering, Mar. 9, 2016, Ferdowsi University of Mashhad, Mashhad, Iran

ISI papers :

- **A Naeim abadi**, F Hemmati, A new model for rubber-toughened polyamide 12 nanocomposite using convolutional neural network methodology, Biomedical image processing J, Under revision