GAP SOLUTIONS
a System One company

Submitted to Amentum for ARPA-H STATS

NAME	JOB TITLE	PWS Tasks
Yuqing Mao	ITDI - Basic Software Developer	7.1.1-7.1.2; 7.1.11-7.1.13

Job Responsibility	Qualifying Skills
Enterprise Software Development & AI Integration: Design, develop, test, and deploy AI-driven enterprise applications.	· Developed machine learning pipelines for AI-driven biomedical applications at National Library of Medicine (NLM), improving medical terminology classification. · Built a state-of-the-art NLP system to extract clinical information from medical notes, converting free-text data into structured formats for regulatory and research applications. · Designed AI-powered text mining algorithms for PubMed, improving biomedical literature searchability.
Cloud Computing & Generative AI Model Development: Implement cloud-based solutions and develop generative AI models to support advanced computing tasks.	· Implemented deep learning models using cloud-based AI solutions, including AWS and Hugging Face. · Developed and optimized generative AI models for biomedical information extraction, recognized as a top performer in BioASQ, an international AI challenge. · Enhanced biomedical NLP pipelines using TensorFlow, integrating AI-based analytics for healthcare applications.
Cross-Functional Collaboration & System Optimization: Work with diverse teams to optimize machine learning algorithms and enhance software performance.	· Worked with interdisciplinary teams at NIH, FDA, and academia to develop AI-based regulatory science tools. · Optimized data processing algorithms for medical records analysis, improving accuracy and reducing computational costs. · Led a team of Ph.D. interns in deep learning model development for medical terminology standardization, achieving a 23% improvement in classification accuracy.

Education

Ph.D., Computer Science, Nanyang Technological University, Singapore   **May 2012**
Dissertation: Epistemology-based Social Search for Exploratory Information Seeking

M.S. and B.S. degrees in Computer Science focusing on Natural Language Processing, Artificial Intelligence, Machine Learning, Information Retrieval

Experience

Research Scientist, National Library of Medicine (NLM), USA   **2018 - Present**

Focus on utilizing deep learning techniques including knowledge graph and generative AI to improve the information extraction and normalization to standard terminologies, for the development of new medicines, treatment pathways, and better patient outcomes.

· Award
  Recipient of NLM Staff Appreciation Award for outstanding contributions to the advancement of Generative AI capabilities across NLM programs.

Research and Development

· Developed and mentored intern Ph.D. students to customize deep learning models for Unified Medical Language System (UMLS) editing. Assigning new atoms to millions of UMLS concepts is a time-consuming and costly task. The automated assistance improved the accuracy from 70% to 93%, which greatly alleviated the burden on UMLS editors and reduced the risk of errors.
· Built a system based on large language models trained on real-world EMR data such as doctor's discharge summaries, to identify and label entities in the clinical notes that correspond to specific medical concepts. It converts the free-text data into a structured format that can be analyzed by computers and is the best US system in the global SNOMED CT Entity Linking Challenge.
· Enhanced biomedical natural language processing (NLP) pipelines using TensorFlow and ETL methods to help FDA Pharmacists detect dangerous drug-drug interactions from the drug labels.

Professor, School of Information Technology, Nanjing University of Chinese Medicine, China   2015 - 2018

Conducted research and built software for natural language processing in Traditional Chinese medicine literature using deep neural network frameworks, taught computer science courses.

· Award and Funding
  Recipient of the National Natural Science Foundation of China, the Six Talent Peaks Program and the Program for High-Level Entrepreneurial Talents Introduction.
· Leadership
  Mentored students in College Student Innovation Projects such as using wearable device sensing to monitor basic human behavior and detect irregular activity pattern changes.
· Entrepreneurship
  Created a startup of Traditional Chinese Medicine Healthcare Robots, generated original ideas, conducted the feasibility study, and drafted the business plan.

Postdoctoral Research Fellow, National Center for Biotechnology Information,
National Institute of Health, USA   2012 - 2015

Undertook research in biomedical text processing, information retrieval, and AI/machine learning, developed machine-learning-based methods such as automatically assigning controlled vocabulary terms (MeSH) to articles, for facilitating semantic-based literature retrieval and providing reliable information that can be trusted by common people.

· Awards
  - Recipient of NIH Fellows Award for Research Excellence (FARE) competition recognizing the work of large-scale literature indexing using machine learning that was deemed outstanding based on the scientific merit, originality, experimental design, and the overall quality and presentation.
  - Winner of the Transinsight Award for Semantic Intelligence, for pushing the limits in the area of Semantic Indexing and Question Answering to a new level, demonstrating the overall best performance in the competition funded by the European Commission, and showing an outstanding ability in natural language processing.
· Algorithm Design and Applications
  - Invented a state-of-the-art machine learning algorithm to automate PubMed articles indexing (7000 new articles per day), which was time-consuming (up to 2 or 3 months) and costly (10 dollars per article) for NLM indexers. By automatically predicting keywords of articles, this algorithm significantly reduced the amount of time and the cost.

- Proposed and implemented new methods of the PubMed search engine query logs analysis, which captured users' behavior to better understand user intent. Search engine can provide authentic results according to differences between professional and general users' different search patterns.
· Collaboration
- Organized the International Automatic Gene Ontology (GO) Annotation task. Collaborated with professional GO curators from top organizations such as Cambridge University, Stanford University and Caltech, to conclude the state-of-the-art in automatically GO annotation mining, which is important to researchers and clinicians interested in the genetic and genomic basis to improve human health or agriculture. This work also highlighted and suggested solutions to the remaining technical challenges.

Research Associate, School of Computer Science, Engineering & Mathematics
Flinders University, Australia **2010 - 2011**

Investigated the key issues in online exploratory health information search, developed a prototype system providing users summarized answers and links to sources as epistemologies, shared by others or generated by AI technologies.

· Proposed the epistemology-based search solution to interactive information seeking through novel user interfaces
· Built a system that combines the power of existing specialized health information search engines and the collective wisdom of human users

ADDITIONAL SKILLS & QUALIFICATIONS

· Skilled in C/C++, Java, Python and other programming languages
· Created state-of-the-art machine learning models for Natural Language Processing applications on large-scale corpus (PubMed, UMLS, Clinical Notes)
· Winner of global machine learning competition BioASQ: a challenge on large-scale biomedical semantic indexing and question answering
· Publications at top journals (e.g., IEEE transactions, JAMIA) and top conferences (e.g., EMNLP, ACM CSCW, IUI)
· Editorial Board: PeerJ Computer Science, PLOS ONE, Frontiers in Computer Science, etc.
· Demo Chair: 2014 IEEE ATC 2014 (The 11th IEEE International Conference on Autonomic and Trusted Computing)
· Program Committee Member:
- 2013 International Workshop on Data Mining for Healthcare (held in conjunction with the IEEE International Conference on Healthcare Informatics, ICHI 2013)
- 2013 BCS-HCI 2013 (The 27th International British Computer Society Human Computer Interaction Conference)
- 2009 IEEE ICEBE 2009 (The 6th IEEE International Conference on e-Business Engineering)
· Reviewer: IEEE Access, Artificial Intelligence in Medicine, Bioinformatics, Database (Oxford), JAMIA, Journal of Medical Internet Research, Journal of the Association for Information Science and Technology, etc.
· Certifications: 2002 Object-Oriented Analysis and Design Using the UML, IBM Rational University

Programming Skills:

· Programming Language: C/C++, Java, Python, Perl, R
· Web Development: JSP, JavaScript, C#, PHP, HTML/XML/CSS, AJAX, SVG
· Programming Platform: Windows, Android, Unix, Linux, mixed-platform, AWS, Hugging Face
· Development Environment/Tools: Visual Studio, Qt, Eclipse, Azure AI Studio, Lucene, Terrier
· Database Programming: Oracle, SQL server, MySQL
· Software Framework: Hadoop, Spark, Tensorflow, Pytorch, Keras, LangChain

Research Techniques:

- Information Retrieval: language model, collaborative filtering, recommender systems
- Natural Language Processing: latent semantic indexing, LDA, knowledge graph, RAG
- Machine Learning: deep learning, learning to rank, SVM, CNN, GCN, Transformer, GPT
- Knowledge Discovery and Data Mining: scikit-learn, classification and clustering for text analysis
- Social networking: social reputation and trust models, social network analysis
- Computer Vision: edge detection, mathematical morphology, content-based image retrieval