

Bradley Knight

Senior AI Engineer and Full-Stack Engineer

✉ bradleyjosephknight@gmail.com

☎ +1 830 214 7220

📍 New Braunfels, TX, United States

🌐 <https://www.linkedin.com/in/bradley-knight-275470310/>

👤 PROFILE

I am a Senior AI Engineer and Full-Stack Engineer with more than 12 years of company work experience, deep experience, and strong abilities in various fields such as machine learning, deep learning, data science, speech recognition, natural language processing, generative AI and frontend/backend development.

- Software Engineering
- Machine Learning
- Natural Language Processing
- Data Science
- Generative AI
- Web development
- Chatbot development
- LLM(Falcon, Llama2, Llama3), Langchain, RAG, Pinecone, OpenAI ChatGPT, Text Embedding
- Python: Django, Flask, Tensorflow, Keras, Pytorch, Scikit Learn
- Full-Stack: MERN, MEAN, Django, Laravel+Vue
- Frontend: HTML, CSS/SCSS, TailwindCSS, Ajax, jquery, Javascript frameworks such as Node.js, React js, Angular js, Vue js, Next js,
- Backend: Node, Express js, Python, Django, Flask, PHP, Laravel, Codeigniter, GraphQL
- Database: MySQL, MongoDB, PostgreSQL, MS SQL, SQLite.
- Service: AWS, Netlify, Heroku, Azure, cPanel

My career has been dedicated to pioneering advanced AI solutions and crafting intuitive web applications from inception to deployment within these critical sectors.

Currently, I would like to contribute as a Senior AI and Full-Stack Engineer to a company that develops high-quality products.

📁 PROFESSIONAL EXPERIENCE

2022/01 – 2024/04
Foster City, California,
United States

Sr. AI Engineer and CV Team Lead

SocialBook

- Led the development of cutting-edge natural language processing (NLP) solutions for various applications, including sentiment analysis, named entity recognition, and text summarization, utilizing Python and libraries such as NLTK and spaCy.
- Developed cutting-edge AI solutions leveraging Large Language Models (LLMs) and Generative AI techniques to transform healthcare processes, including automated document processing and customer service chatbots.
- Spearheaded the implementation of state-of-the-art Large Language Models (LLMs) such as GPT and BERT for natural language understanding and generation tasks, utilizing frameworks like Hugging Face Transformers and TensorFlow.
- Developed innovative solutions using Langchain, a novel framework for language modeling and text generation, enhancing the capabilities of conversational AI systems and content generation pipelines.
- Designed and implemented Retriever-Answerer-Generator (RAG) models for question answering and document summarization tasks, leveraging pre-trained language representations and efficient retrieval mechanisms.

- Deployed AI-powered search solutions using Pinecone, a vector database for similarity search, enabling fast and scalable retrieval of semantically related documents and entities.
- Pioneered the development of generative AI models for creative applications such as image generation, content writing, code generation, video generation, lead generation, text-to-image, and image-to-image synthesis, as well as music composition, utilizing Python and frameworks like TensorFlow and PyTorch.
- Developed and deployed NLP solutions for various business use cases, including chatbots, virtual assistants, and information retrieval systems, using Python and frameworks such as Django and Flask.
- Implemented custom NLP pipelines for data preprocessing, feature extraction, and model training, optimizing performance and scalability.
- Developed an MLOps pipeline for training, deploying, monitoring, and managing machine learning models in production environment with high availability features enabled using Kubeflow, TensorFlow Extended Airflow, MLFlow, etc

2019/10 – 2021/06
Ohio, United States

Artificial Intelligence Engineer

Vpl Solutions, LLC

- Developing robust software for integrating multiple sensors and object tracking systems.
- Develop novel real-time 3D scene reconstruction techniques and delivering accurate visual odometry systems
- Developed scalable and secure software solutions for financial institutions using Java, Python, and Hadoop, focusing on fraud detection and risk management systems.
- Spearheaded the implementation of state-of-the-art Large Language Models (LLMs) such as GPT and BERT for natural language understanding and generation tasks, utilizing frameworks like Hugging Face Transformers and TensorFlow.
- Developed state-of-the-art natural language processing (NLP) solutions using Python and its associated libraries such as NLTK, spaCy, and Gensim.
- Implemented advanced deep learning models including recurrent neural networks (RNNs), long short-term memory (LSTM) networks, and transformer architectures like BERT and GPT, using TensorFlow and PyTorch frameworks.
- Deployed AI-powered search solutions using Pinecone, a vector database for similarity search, enabling fast and scalable retrieval of semantically related documents and entities.
- Decision Trees, Clustering, Customize Neural Networks, and SVMs were used to analyze customer data such as purchasing power, demographics, and other behavioral data to identify customer personality traits.
- Led the integration of NLP technologies into production systems using Docker and Kubernetes for containerization and orchestration, ensuring scalability and reliability.

2015/01 – 2019/05
Baltimore, Maryland,
United States

Full Stack Engineer

RedHorizonPLC

- Contributed to the development of software applications, including web and mobile platforms, using Python, Java, and JavaScript.
- Developed web and mobile applications using full-stack technologies such as Python/Django, JavaScript/React.js, and Java/Android SDK.
- Contributed to the development of full-stack web applications, including frontend UI design, backend server logic, and database management, using technologies like HTML/CSS, JavaScript, Python, and Django.
- Developed scalable and maintainable web applications for generative AI services using TensorFlow, PyTorch, Django, Flask, Node.js, React.js, Next.js and MongoDB.
- Implemented RESTful APIs for seamless communication between frontend and backend systems, enabling efficient data exchange and integration with third-party services.

SKILLS

AI Development

Machine Learning, DeepLearning, Natural Language Processing, Computer Vision, Data Science, Generative AI, LLM, ChatGPT, Langchain, RAG, Pinecone, Object detection, Object tracking, Image segmentation, Speech recognition, Sentiment Analysis, Tensorflow, Pytorch, Keras, Scikit-Learn, CNN, RNN, Mask R-CNN, OpenCV, YOLO, GANs, Stable Diffusion, OCR

Full-Stack Development

Python, JavaScript, Node.js, C++, C#, Java, React.js, AngularJS, Next.js, Vue.js, Three.js, TypeScript, Django, Flask, Express.js, REST API, Fast API, API Integration, MySQL, MongoDB, PostgreSQL, SQLite

Cloud Engineering and DevOps

Docker, Kubernetes, Git, AWS, GCP, Netlify, Heroku, Azure, cPanel

LANGUAGES

English



EDUCATION

2011/08 – 2015/04

Bachelor of Computer Science (BCompSc)

Austin, Texas,

The University of Texas at Austin

United States