

# Karthek Manda

Toronto, Canada — (519)-984-4855

Karthek3197@gmail.com — <https://linkedin.com/in/karthekmanda> — <https://github.com/karthek3197> — <https://karthek3197.github.io/>

---

## SUMMARY

**Data Scientist (Machine Learning) with over 4 years of experience** in NLP, Machine Learning, and Demand Forecasting. Expertise in developing AI-driven solutions like RAG models, NLP tools, and time-series forecasting, improving data quality, efficiency and customer satisfaction. Skilled in models like **BERT**, **GPT-4**, **LLaMA2** and **ARIMA**, with hands-on experience in **LangChain**, **TensorFlow** and **AWS SageMaker**. Strong ability to **build models from scratch** with deep **theoretical knowledge** of various ML Techniques and statistical analysis. Proven success to automate processes, reduce manual intervention, and deliver real-time insights for high-impact applications.

---

## EXPERIENCE

### Data Scientist - ML

Nov 2022 – Present

**Riskthinking.ai** - <https://riskthinking.ai/> - Toronto, Canada

- Developed and deployed the **ask-riskthinking.ai** chatbot using LLaMA2 for natural language understanding, and utilized LangChain to orchestrate interactions between the chatbot and internal databases, enabling accurate and real-time responses to climate risk-related queries.
- Developed and deployed a **Retrieval-Augmented Generation (RAG)** model using BERT and GPT-4o to automate the extraction of key parameters from sustainability reports, improving data quality for climate risk assessments by 10% and eliminating manual intervention.
- Developed an **NLP-based company matching tool** using cosine similarity and metadata transformation, increasing data processing speed by 50%, resulting in more reliable climate risk analytics across datasets.
- Built a **value-chain analysis tool** using **GPT-3.5** and **LLaMA2** to identify competitors, suppliers, and customers, delivering strategic insights into company dependencies and enhancing the platform's value in risk assessments for clients.
- Deployed machine learning models at scale using **Docker** on **AWS-SageMaker**, ensuring scalability, reliability, and high availability of machine learning solutions.

### Machine Learning Scientist

Nov 2021 – Nov 2022

**Okos Smart Homes** - <https://okos.ca/index.html> - Calgary, Canada

- Developed **demand forecasting models** for energy consumption in smart homes using **time-series** techniques (ARIMA, Attention models). Enabled smart systems to optimize energy usage, leading to a 15% reduction in energy costs and improved customer satisfaction.
- Developed AI-backed smart home solutions, including predictive maintenance and energy optimization systems, leading to a 15% reduction in energy costs for customers. This also resulted in a 20% increase in **customer retention**.
- Developed personalized customer behavior prediction models using **LSTM** and **Transformer-based models** to analyze energy consumption patterns. This model improved energy efficiency, leading to a 15% increase in user engagement as customers interacted more frequently with the system to optimize their energy use.

### Junior Data Scientist

Dec 2018 – Dec 2019

**IPS Engineering Consultants** - <https://ipsengineers.com/> - India

- Developed NLP techniques (**NER**, **text classification**) to automate key information extraction from technical documents, reducing manual processing by 40% and improving data accuracy and workflow efficiency.
- Created and maintained detailed **technical documentation** for model development, testing, and deployment, improving team collaboration and streamlining onboarding for new members.

---

## TECHNICAL SKILLS

- **Programming Languages:** Python(OOP, Design Patterns), SQL, R
- **Machine Learning & Deep Learning:** Supervised Learning, Unsupervised Learning, feature Engineering, Self-supervised learning, Deep Learning, Reinforcement Learning (novice)
- **Data Manipulation & Statistics:** NumPy, Pandas, SciPy, PCA, TSNE, A/B testing, K-S test, Chi-Square Test, Z-score

- **Frameworks:** Hugging Face, TensorFlow, Keras, Scikit-Learn, Flask
  - **Databases and Version Control:** MySQL, PostgreSQL, Git
  - **Cloud Platforms:** AWS (SageMaker), GCP (Vertex AI)
  - **Natural Language Processing (NLP):** NLTK, SpaCy, NetworkX, LangChain
  - **Data Visualization:** Matplotlib, Seaborn, Plotly
  - **DevOps & Deployment Tools:** Kubernetes, Docker, System Design (naive)
- 

## EDUCATION

**Masters - Computer Engineering**  
*University Of Windsor - Windsor, Canada*

**Jan 2020 – Apr 2021**