Blend ⊖

AI Data Scientist - Manager

Columbia, MD

Full-time

Company Description

Blend360 is a world class marketing, analytics, and technology company that delivers the best results for our clients. Our primary focus is Data Sciences; leveraging data and applied mathematics to solve our clients' business challenges. Blend360 is known for our exceptional people, our get-it-done mentality, and delivering high impact and sustainable results. If you love to solve difficult problems and deliver results; if you like to learn new things and apply innovative, state-of-the-art methodology, join us at Blend360.

Job Description

- You will work as part of our global Data Science team to provide data driven AI solutions for our customers using state-of-the-art methods and tools.
- Work with practice leaders and clients to understand business problems, industry context, data sources, potential risks, and constraints
- Work with practice leaders to get stakeholder feedback, get alignment on approaches, deliverables, and roadmaps
- Create and maintain efficient data pipelines, often within clients' architecture. Typically, data are from a wide variety of sources, internal and external, and manipulated using SQL, spark, and Cloud big data technologies
- Assemble large, complex data sets from client and external sources that meet functional business requirements.
- Build analytics tools to provide actionable insights into customer acquisition, operational efficiency, and other key business performance metrics.
- Perform data cleaning/hygiene, data QC, and integrate data from both client internal and external data sources on Advanced Data Science Platform. Be able to summarize and describe data and data issues
- Utilize deep learning principles and architectures, including CNNs, RNNs, and transformers, apply these techniques to natural language processing tasks.
- Manipulate model parameters to achieve desired outcomes in text generation.
- Craft effective prompts that guide AI models to generate desired outputs. Understand how different prompt structures influence AI behavior.
- Use RAG models, and combine a retrieval component with a generator to enhance the quality and relevance of the AI's output. Understand how to effectively integrate external knowledge sources into AI responses.
- Train and fine-tune models on specific datasets to improve performance and ensure the relevance of the outputs to the task at hand.
- Conduct statistical data analysis, including exploratory data analysis, data mining, and document key insights and findings toward decision making
- Document predictive models/machine learning results that can be incorporated into client-deliverable documentation
- Assist client to deploy models and algorithms within their own architecture

Qualifications

- Profound knowledge of deep learning principles and architectures, including CNNs, RNNs, and transformers, with the

ability to apply these techniques to natural language processing tasks.
- Deployment experience - Understanding of how to integrate models into production
- Engineering experience
- In-depth understanding of the workings of LLMs and the ability to manipulate model parameters to achieve desired outcomes in text generation.
- Expertise in crafting effective prompts that guide AI models to generate desired outputs. Understand how different prompt structures influence AI behavior.
- Experience with RAG models, which combine a retrieval component with a generator to enhance the quality and relevance of the AI's output. Understand how to effectively integrate external knowledge sources into AI responses.
- Capability to train and fine-tune models on specific datasets to improve performance and ensure the relevance of the outputs to the task at hand.

- MS degree in Statistics, Math, Data Analytics, or a related quantitative field
- At least 3 years of post graduate professional experience in Advanced Data Science, such as predictive modeling, statistical analysis, machine learning, text mining, geospatial analytics, time series forecasting, optimization
- Demonstrated Experience with NLP and other components of AI
- Experience implementing AI solutions
- Experience with one or more Advanced Data Science software languages (Python, R, SAS)
- Proven ability to deploy machine learning models from the research environment (Jupyter Notebooks) to production via procedural or pipeline approaches
- Experience with SQL and relational databases, query authoring and tuning as well as working familiarity with a variety of databases including Hadoop/Hive
- Experience with spark and data-frames in PySpark or Scala
- Strong problem-solving skills; ability to pivot complex data to answer business questions. Proven ability to visualize data for influencing.
- Comfortable with cloud-based platforms (AWS, Azure, Google)
- Experience with Google Analytics, Adobe Analytics, Optimizely a plus