SAP Summer Internship Program – SAP Ariba Data Science

SAP Ariba is the marketplace for modern business, creating frictionless exchanges between millions of buyers and suppliers across the entire source-to-pay process. Our market-leading solutions enable companies to simplify collaboration with their trading partners, make smarter business decisions and extend their collaborative business processes with an open technology platform. More than two million companies use SAP Ariba solutions to connect and collaborate around nearly one trillion in commerce on an annual basis.

As a Data Science Intern, you will use data science to create intelligent products for the best-in-class cloud procurement solution. You will be a part of our Business Network team and help draw insights from one of the biggest B2B data stores in the world. A key trait to be successful in this role is to have a love for data and utilize it to solve critical problems and deliver end user value. You will create predictive models, extract insights from data, find patterns in payments and buying data, and use your analytical skills to take our business applications to the next level.

Key Responsibilities:

- Work with other data scientists, developers and cross functionally with product managers and other engineering teams to deliver predictive models from concept to product.
- Model Building Cycle – pull, cleanse, and validate data for analysis and modeling.
- Create and implement predictive models for various business process like payment risk, relationship discovery, categorization (unsupervised learning), catalog optimization, NLP, etc.
- Work with software engineers to put predictive models in production
- Determine the tracking necessary to enable analytics of our products and features by working closely with product and engineering partners

Key Qualifications:

- MS or PhD Candidate in statistics, economics, physics, mathematics, computer science, or related discipline
- Strong understanding of applied statistical concepts such as regression, time series analysis, machine learning (supervised & unsupervised learning such as segmentation, classification and clustering)
- Hands-on experience building models and machine learning algorithms
- Experience coding in Python, Java, C, C++ and/or SQL a plus
- Knowledge of or experience working with big data (e.g., Hadoop) a plus
- Be an iterative and quick thinker – you deliver preliminary results fast and then iterate over time