**Nolans Input:**

Koushika leverages the Azure Databricks platform to code in R, Python, and Scala. In addition to API connections for applications, data collection, and cloud resources she is required to write machine learning and AI algorithms for the credit union. Specifically we leverage these models to solve advanced business questions. Examples of this are:

* **K-means clustering:** Segmentation of members based on preferences and behaviors so that we can specialize their experience to fit them.
* **Logistic Regression testing:** A good method for the credit union to test predictions of decision based problems.
* **Linear Regression:** Determining the variables that lead to member churn rate and creating predictive equations
* **Forecasting:** Leveraging historical data to predict future values of important financial values like production, loss, and economic indicators.
* **Text Analytics:** Using transcribed audio files to gain both sentiment and entity analysis of calls that come in for root cause analysis

Koushika leverages her expertise to develop and maintain the credit union’s algorithms while also serving as an expert consultant on all advanced statistical analysis. The development and upkeep of these models takes approximately 80% of her time with the remaining 20% being spent on consultancy.

As a part of the team, Koushika has been an integral part in the development of an in house risk mitigation model. Her skills developed from her work experience in India and studies from the University of South Florida (USF) have proved invaluable in this endeavor. This project being one of many on the list that Suncoast Credit Union has on the horizon. Having Koushika as part of the team allows us to engage in rapid development, iteration, and maintenance of these powerful models.

[9:51 AM] NOLAN WALKER III

This is an expansion of your job duties which are listed as this:

* Works with datasets to create equations based on advanced statistics and mathematical results.
* Identifies action from predictive and prescriptive results.
* Tracks trends in data and measures success of actions.
* Designs analysis plans, conducts advanced statistical analysis, produces results, helps with implementation of equations to segment members (or transactions) and monitors populations to quantify results. Ideally this process leads to recurring best practices.
* Responsible for custom Suncoast Credit Union analytics, including the SunLogix, solution. This includes creating, maintaining, and troubleshooting models, lift charts, etc
* Writes, reviews, and maintains documentation describing all elements of data warehouse and cube design (including standards, functionality, logic, deployment and implementation procedures, time estimates, and test plans).
* Maintains knowledge and understanding of current trends, laws, and issues affecting area of expertise. Attends educational events that will increase professional knowledge and be otherwise beneficial to the Credit Union. Completes annual BSA/AML Compliance Training and understands employee’s role in maintaining an effective BSA/AML compliance program, and completes FACT Act Red Flag training.

**Koushika’s Edit:**

**JOB DESCRIPTION**

Job Title: Data Scientist

Qualifications and Skills:

* Bachelor / Master’s degree in Business Analytics, Statistics, or related field
* Extensive background in data mining and statistical analysis
* Excellent skills in predictive modeling and machine learning skills
* Experience with programming languages such as R, Python, Scala
* Experience with relational databases and SQL scripting
* Excellent Data cleaning, wrangling and feature engineering skills
* Developing data science pipelines.

ROLE MAPPING:

* Works with datasets to create equations based on advanced statistics and mathematical results.
* Identifies action from predictive and prescriptive results.
* Tracks trends in data and measures success of actions.
* Designs analysis plans, conducts advanced statistical analysis, produces results, helps with implementation of equations to segment members (or transactions) and monitors populations to quantify results. Ideally this process leads to recurring best practices.
* Responsible for custom Suncoast Credit Union analytics, including the SunLogix, solution. This includes creating, maintaining, and troubleshooting models, lift charts, etc
* Writes, reviews, and maintains documentation describing all elements of data warehouse and cube design (including standards, functionality, logic, deployment and implementation procedures, time estimates, and test plans).
* Maintains knowledge and understanding of current trends, laws, and issues affecting area of expertise. Attends educational events that will increase professional knowledge and be otherwise beneficial to the Credit Union. Completes annual BSA/AML Compliance Training and understands employee’s role in maintaining an effective BSA/AML compliance program, and completes FACT Act Red Flag training.

**Expanded Job duties:**

Data Design (15%)

* + Formulate datasets required for the business problem from different data sources using relational databases and SQL.
  + Creatively synthesize internal data with the external data sources (such as - demographic data, economic data, etc.) to enhance / enrich the existing data to support further data analysis.
  + Assess the effectiveness of new data sources
  + Extract data from SQL and perform data cleaning, transformation and feature engineering to prepare data for analysis and further modeling.
* Responsible for collecting, cleaning and munging of data to meet the company’s purpose.
* Acquire data, process and clean the data. Integrate and store the data for data science project purposes.

Data Analysis and Visualization (15%)

* + Partner with internal data engineering team to ensure data is captured in analytical environments and establish the data pipeline
  + Initial data investigation and exploratory data analysis
  + Apply advanced statistical approaches and data visualization techniques to capture key trends in the data and derive actionable insights.
  + Interpret outcome of a statistical test and translate it to the business unit in the form of actionable insights.
  + Capturing the key trends from the analysis using Power BI
  + Data discovery and pattern recognition, anomaly detection and run diagnostics.
* Perform initial data investigation and exploratory data analysis.

Model Development, Testing and Maintenance (40%)

* Business solutions, POCs and R&D work
* Work in Azure Cloud platform to develop data science solution using Azure Databricks, DataLake with knowledge in statistical programming using R, Python and Scala.
* Research, study and investigate new relevant Data Science technologies and demonstrate the value and usability with proof-of-concepts.
* Developing data science solutions on big-data using open source tools and cloud computing platforms.
* Building machine learning models through all phases of development from design through formulation, training, evaluation, validation and implementation.

Model Documentation and Explainability (10%)

* Document all processes and research.

Communicating Insights with management as well as other business units (10%)

* Partner with business to identify opportunities to develop data-driven solutions to improve business outcomes.
* Connect with stakeholders to gain a full understanding of the problems they’re looking to solve.
* Presenting findings and solution to various levels of management and key decision makers
* Ability to articulate trends and potential clearly and confidently
* Ability to translate and tailor technical findings into business applicable material with clear recommendations and insights relevant to the audience at hand.
* The reports and presentations will not only be translation of technical analysis into business language but also be simple, concise and understandable and convincing – which requires good communication skills

Compliance and Knowledge Update (10%)

* Adhering with the industry compliance and norms
* Continually research and evaluate emerging technologies in the data science domain as well as the industry.
* Stay current on published state-of-art methods, technologies, and applications and seek out opportunities to apply them.

**General points:**

* Create experimental frameworks for product development and machine learning with the aim to lay strong data foundation for robust analytics to be performed.
* Apply data science/ machine learning models that best fits the problem in hand ( can vary from regression, linear, non-linear, classification, and categorization models in supervised and unsupervised algorithms)
* Measure and improve the results of the mathematically trained machine learning model. Make adjustments based on feedback.
* Track the model performance and sanity check on the existing models and re-train the models as required by the drift categorization.
* Provide model explainability and document the model, underlying logic using mathematics and explaining the model performance in terms of measurable metrics and error estimates.
* Create model accountability by developing processes that aid in model’s decision making and transparency.
* Design data modeling processes to create algorithms and predictive models and perform custom analysis.
* Data analysis: Manipulate large and complex datasets and use them to identify trends and reach meaningful conclusions to inform strategic business decisions.
* Data engineers: clean, aggregate and organize data from disparate sources and transfer it to data warehouses.
* Communicate insights across a diverse audience across all levels of an organization
* Implement algorithms and statistical models to enable computer to learn from the data and make decisions without human intervention.
* Conduct analytical experiments and evaluate theoretical alternate models.

\*\* FROM LINKEDIN\*\*

* Communicate analytic solution to stakeholders and implement improvements as needed to operational systems.
* Implement analytical models into production by collaborating with software developers and machine learning engineers
* Analyze data for trends and patterns. Interpret data with a clear objective in mind.
* Devise and utilize algorithms and models to mine big data stores, perform data and error analysis to improve models, and clean and validate data for uniformity and accuracy.
* Identify relevant data sources and sets to mine for specific business solution / project. And collect large structured and unstructured datasets and variables.
* Execute analytical experiments methodically to help solve business problems and make a true impact across various domains and industries.

**Projects worked on:**

Work samples:

