PROLEGO

Data Scientist Career Map

What's the purpose of this table?

- Provide clearly defined expectations for those at various levels of the organization.
- Provide clear targets for how one can advance/be promoted along a technical track. (Management will be a parallel track.)
- See this is a "playbook" for understanding what various technical roles at Prolego should look like.

There are four categories of skills that are equally important. Everyone will bring a different set of skills to the table, but these categories attempt to capture generally applicable skills.

01 Data and modeling skills
This category addresses technical skill, including specific and general data and ML techniques, technical problem solving, and the ability to tackle challenges at varying levels of complexity.

02 Software skills
This category covers coding ability, including general development best practices, version control, testing, reproducibility, etc.

03 Organizational skills
This category addresses how someone interacts with the rest of the organization, including things like work planning and scoping, risk identification and mitigation, the amount of oversight needed or given, client interactions, and contribution to other organizational efforts like sales and marketing.

04 Professional skills
This category speaks to general professionality, both internally and with clients.

The descriptions of skills within each of these categories is intended to be actionable (what you should be doing to be successful at that level) and as unambiguous as possible.

LEVEL	DATA AND MODELING SKILLS	SOFTWARE SKILLS	ORGANIZATIONAL SKILLS	PROFESSIONAL SKILLS
Data Scientist	• Independently ingests, cleans and structures labeled data • Engineers features appropriate for traditional ML models • Independently performs exploratory data analysis • Applies a wide variety of traditional machine learning models, understanding the requirements, constraints and trade-offs of each • Troubleshoots, diagnoses and remedies basic model development issues	• Writes Python code using a combination of notebooks, scripts and modules • Effectively uses basic DS/ML Python libraries such as pandas, numpy, matplotlib and sklearn. • Includes object-oriented features in code where appropriate • Uses version control on code bases	• Executes tasks in a generally independent manner, while recognizing when to seek guidance and feedback • Proactively communicates risks and challenges that emerge in the course of work • Is an active participant in meetings with client data scientists	• Exhibits general professional skills such as punctuality, clear and timely communication, and takes ownership over own work • Competently and professionally completes technical tasks assigned by more experienced staff • Accepts and responds to feedback and sees every project as a learning opportunity • Demonstrates time management skills and the ability to balance various demands
Senior Data Scientist	• Ingests and structures a wide variety of data (structured, unstructured, noisy, etc.) • Applies a wide variety of ML models, including neural networks and other non-sklearn models • Troubleshoots, diagnoses and remedies most model development issues • Identifies and applies new techniques to solve challenges	• Consistently writes clean, well-structured code that adheres to best practices like version control, testing and reproducibility • Develops moderately complex code projects from scratch • Contributes to internal or external software development teams using a version control system's branching, pull requests, code reviews, integrated ticketing, etc.	• Breaks down and executes a given high-level technical task • Independently executes technical work with minimal oversight • Proactively identifies risks and challenges in upcoming work • Provides reasonable timelines for completion of work • Proactively communicates the status of ongoing work • Communicates the business value of their work • Provides clear written and oral reports to client staff and management	• Develops expertise in one or more specific technical areas • Demonstrates clear and effective communication internally and with clients
Lead Data Scientist	• Frequently modifies, extends or combines ML models to best solve the problem at hand • Trains and develops models to the point of diminishing returns or structural limitations • Proactively seeks out new or emerging techniques from blogs and research papers	• Develops complex code projects from scratch • Applies emerging software requirements like CI/CD, containerization, API calls, etc., where appropriate	• Scopes and plans a technical piece of work given a business objective, providing milestones and reasonable estimates of completion dates • Provides feedback, technical reviews and technical guidance to peers and less experienced staff • Suggests improvements to best practices • Routinely works directly with client personnel to ensure smooth project progress and delivery • Works with external and internal engineering resources to ensure successful solution deployments • Occasionally supports technical sales and marketing efforts	• Demonstrates first-principles thinking in approaching new problems • Displays a project-level sense of ownership • Excellent and proactive communicator internally and with clients
Principal Data Scientist	• Translates complex ideas from research papers into operational code for prototyping, experimentation and/or client use • Identifies and promotes emerging areas or ideas that could benefit present or future clients	• Develops complex code projects from scratch • Applies emerging software requirements like CI/CD, containerization, API calls, etc., where appropriate	• Identifies opportunities for innovation beyond current project scopes • Works with client management and internal managers on topics of feasibility and strategic direction • Is a go-to resource to support technical sales and marketing efforts • Supports ad hoc technical requests from across the organization	• Demonstrates first-principles thinking in approaching new problems • Displays a project-level sense of ownership • Excellent and proactive communicator internally and with clients