

Data Scientist - Data Quality and Validation

Position Overview

The Raynor Cerebellum Project (RCP) aims to develop the next generation of closed loop neuromodulation. To this end, the RCP conducts multi-site clinical research involving the collection of human neurophysiological and behavioral data in Epilepsy Monitoring Unit (EMU) and Neural Monitoring Unit (NMU) sites. We are seeking a data scientist to lead the technical quality control (QC) function that determines whether data collected at partner sites meets defined scientific and technical standards for use in downstream research and modeling. The Data Scientist will also have the opportunity to explore this one of a kind human dataset in collaboration with RCP Institute members.

The position carries substantial independence and technical responsibility. The successful candidate will be expected to design, implement, and maintain the QC methodology and supporting infrastructure directly, including analysis pipelines applied to raw data collected using intracranial recording and stimulation systems. This is not a role in which an established framework will be provided; rather, the candidate will be responsible for developing appropriate methods and tooling in consultation with data collectors and other program stakeholders. The data scientist's exploration of this unique dataset could uncover biological insights critical to accelerating the development of next-generation closed-loop neuromodulation strategies.

Responsibilities

- Determine, on the basis of defined and objective technical criteria, whether incoming data sets from participating sites meet the quality standards necessary for use in downstream research and modeling.
- Develop and deploy formalized metrics to assess the quality of neural and behavioral data collected at partner sites, including electrophysiological recordings and associated behavioral task data.
- Design and apply model-based validation approaches to assess whether collected data support their intended analytic and scientific uses.
- Apply machine learning methods, as appropriate, to support anomaly detection and to improve the efficiency and consistency of quality control procedures as data volume increases across sites.
- Produce data determinations supported by clear, objective, and reproducible rationale.
- Maintain records of quality control outcomes across sites over time, and report on patterns relevant to site performance, training needs, or contractual considerations.
- Collaborate with site personnel and other program stakeholders on troubleshooting and resolution of data quality issues.
- Contribute to the ongoing development and refinement of quality control standards and methodology as the program evolves.
- Collaborate with Institute members to explore the utility of the data for the RCP mission.

Qualifications

- PhD in Computational, Systems, or Cognitive Neuroscience or related field.
- Experience in the analysis of sEEG data or comparable intracranial or high-density electrophysiological data (e.g., ECoG, single- or multi-unit recordings), including relevant signal processing and quality assessment methodology.

- Experience analyzing human behavioral data, including three-dimensional kinematic data (e.g., motion capture, pose estimation, or movement trajectory data).
- Applied machine learning experience specific to decoding or classification of neural time-series data (binary, multi-class, or multi-label).
- Experience with spectral and time-domain feature extraction from electrophysiological signals (e.g., band power estimation, short-time Fourier transform or spectrogram-based representations, and time-domain approaches).
- Familiarity with spatial referencing schemes used in intracranial or scalp electrophysiology (e.g., Laplacian referencing, common average referencing) and the considerations governing their appropriate use.
- Demonstrated capacity to work independently, including the ability to design and implement analytical systems and infrastructure without reliance on an existing framework.
- Strong proficiency in Python (or an equivalent language) for scientific computing, signal processing, and pipeline development.
- Sound scientific judgment in translating heterogeneous, real-world clinical data into objective and reproducible quality determinations.
- Strong written communication skills, with the ability to document technical rationale clearly for both scientific and non-technical audiences, including operational and contractual stakeholders.

Preferred Experience

- Experience with data collected in clinical or hospital-adjacent settings (e.g., epilepsy monitoring units, neural monitoring units, or comparable environments).
- Familiarity with neural recording or stimulation hardware (e.g., Ripple Neuromed, Blackrock, Tucker-Davis, or comparable systems).
- Experience coordinating data quality efforts across multiple collection sites with varying equipment, protocols, or personnel.
- Familiarity with fine-tuning pretrained or foundation models on downstream decoding tasks (i.e., transfer learning applied to neural or physiological time-series data).
- Familiarity with benchmark frameworks for neural decoding (e.g., Neuroprobe or comparable frameworks) and associated evaluation conventions, including temporal train/validation/test splitting.
- Familiarity with HIPAA-compliant data handling and de-identification standards.

Position Summary

This role serves as the principal technical authority on data quality within the program, providing the analytical judgment on which site partners, internal stakeholders, and downstream modeling efforts rely. The candidate is expected to develop quality control methodology and infrastructure that scales in rigor and efficiency as the number of participating sites and the volume of collected data increase.

Location: Position will be in-person based with RCP teams in Fort Worth, TX; Durham/Chapel Hill, NC; or Denver, CO.

Pay Range: \$120-200k/year

To Apply: Submit a cover letter and resume to: careers@onceuponatime.org.