

データエンジニア Biology Machine Learning

BI・DWHエンジニアのご経験のある方は歓迎です。

募集職種

人材紹介会社
株式会社ジェイ エイ シー リクルートメント

採用企業名
非公開

求人ID
1624703

業種
医薬品

雇用形態
正社員

勤務地
東京都 23区

給与
600万円 ~ 1200万円

勤務時間
09:00 ~ 18:00

休日・休暇
詳細は求人ご紹介時にご案内いたします。

更新日
2026年11月26日 16:11

応募必要条件

キャリアレベル
中途経験者レベル

英語レベル
流暢レベル

日本語レベル
ネイティブレベル

最終学歴
大学卒： 学士号

現在のビザ
日本での就労許可が必要です

募集要項

【求人No NJB2426509】
【具体的な業務内容】
1. データパイプラインの構築
■株間やエンジニアリングされた変異株における遺伝子型の差異（SNP、インデル、遺伝子挿入/欠失、プラスミド構成など）を追跡し、各表現型（フェノタイプ）がそれを生み出した正確な遺伝子型と紐づく状態を維持する。
■社内アッセイから得られる表現型ラベル（目的物質の生産量・力価、酵素活性、増殖速度など）を収集・キュレーションする。
■オミクスラベル（トランスクリプトミクス、メタボロミクス）を一貫した処理、メタデータ管理、品質チェックのもとで収集・キュレーションし、ゲノムデータと紐づける。
■DNA、タンパク質、全ゲノムタスクに関する社内ベンチマークスイートを拡張する。

2. ラボ研究員との協働
■ラボ研究員と連携し、サンプル、菌株、プレート、アッセイ結果の命名規則・記録方法・紐づけルールを定義し、あらゆるデータポイントが元の実験まで完全に追跡（トレーサビリティを確保）できるようにする。
■ウェットラボ実験用のメタデータ標準規格およびデータテンプレートを設計し、ラボでのLIMS（ラボ情報管理システム）への定着を支援する。
■研究員がエンジニアの手を借りずに自らデータをアップロード・検証・探索できるシンプルなツールを開発する。

3. インフラストラクチャとデータ品質の担保
■LIMSとデータプラットフォーム間の連携を担い、サンプル、プレート、菌株、アッセイ結果がラボから学習データへとエンドツーエンドでID追跡された状態で流れる基盤を構築・維持する。
■バージョン管理されたインプット、固定されたツール環境（Pinned Tools）、データカード、生ファイルから学習用データセットまでのリネージ（来歴管理）を通じて、すべてのデータセットの再現性を担保する。
■クラウド上のGPU/CPUクラスターにおいて、最適化されたワークロードを実行・運用する。

【レポートライン】Principal AI Scientist（US CA州在住）
※レポートラインのUSチームとは英語 / 日本のチームメンバーとは日本語を使ってのコミュニケーションになります

■ミッション
機械学習モデルを支える「データ層」の全体設計・運用を担っていただきます。 ゲノム側では、天然株やエンジニアリングされた変異株全体の遺伝子型（ジェノタイプ）変化を追跡し、ラボ側では、bit EVOのDBTLサイクルから生み出されるウェットラボデータ（シーケンスリード、プレートアッセイ、オミクスデータなど）をモデル学習に適したラベルへと変換します。
配属先は、研究者と計算科学者（コンピュテーショナルサイエンティスト）で構成される少人数で意思決定の早いチームです。あなたが構築するデータ基盤とワークフローこそが、当社のAIモデルが学習するための源泉となります。

スキル・資格

【必須スキル・経験】
・企業または研究機関におけるデータパイプライン構築の実務経験（2年以上）
・高いPythonスキル、実践的なSQLスキル、Linuxコマンドライン操作の習熟
・バイオインフォマティクスの主要ファイル形式（FASTA、GFF/GTF、GenBank、シーケンス/アライメント出力形式など）の実務的な理解
・クラウドのオブジェクトストレージおよびコンピューティング（AWS、GCP、Nebius等）、Docker、Gitの使用経験

■語学
・日本語　ビジネスレベル以上　/ 英語　ビジネスレベル以上以上

会社説明

ご紹介時にご案内いたします