データエンジニア / バイオロジー×機械学習

唯一無二のバイオビッグデータで、次世代の生物言語モデルを拓くポジション

募集職種

採用企業名
bitBiome株式会社

求人ID
1621012

部署名
R&D

業種
その他

会社の種類
中小企業 (従業員300名以下)

外国人の割合
外国人 少数

雇用形態
正社員

勤務地
東京都 23区, 新宿区

最寄駅
東西線、　早稲田駅

給与
600万円 ~ 1200万円

勤務時間
基本9：00 ~ 18：00 / 専門業務型裁量労働制：1日の労働時間 8時間とみなす

休日・休暇
完全週休2日制

更新日
2026年11月23日 00:00

応募必要条件

職務経験
3年以上

キャリアレベル
中途経験者レベル

英語レベル
ビジネス会話レベル (英語使用比率: 50%程度)

日本語レベル
ビジネス会話レベル

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

最終学歴
大学院卒： 修士号/博士号

現在のビザ
日本での就労許可が必要です

募集要項

≪募集要項・本ポジションの魅力≫

- 世界でも唯一無二の「超巨大・独自データ（35億遺伝子）」を扱うことができます。
- IT企業では体験できない「生物実験×AI」/「自分が整えたデータとモデルによって、現実世界で新しい酵素や菌株が生まれ、機能する」というリアルワールドへのインパクトを実感できます。
- 単なるデータパイプライン構築に留まらず、AIモデリングへも関与できます。
- 少人数×グローバル水準のハイレベルなチーム環境です。

【ミッション】

当社の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。

その後、これら双方のデータを統合し、最新の実験結果が届くたびに更新される、モデル学習・評価用の高品質なデータセットへと昇華させます。

また、モデリング自体の高度化にも深く関与していただきます。ラボの研究員やAIサイエンティストと密に連携しながら、「データのどこがモデル性能のボトルネックになっているか」を見極め、ゲノムやラボ測定値のより良い表現方法を提案し、新たなモデルアーキテクチャの設計・検証にも携わっていただきます。

【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。

【配属先】研究開発部(R&D) Comutational Biology

【レポートライン】Principal AI Scientist（US CA州在住）

【チーム構成】Head of Innovation1名（US在住）、Principal AI Scientist 1名（US在住）、スタッフ2名（日本・本社オフィス勤務）

■当社について：

地球上に存在する微生物の大半は、これまで一度もDNA配列が解読（シーケンス）されていません。早稲田大学発のスタートアップであるbitBiomeは、独自の機械学習、ゲノミクス、そしてウェットラボ（実験環境）をひとつの拠点に統合し、この未知なる微生物の多様性をサステナブルな製品へと変換しています。

■当社の魅力：
・ベンチャー企業ならではの風通しの良さがあります。
・ビジネスの成長と自身の成長をダイナミックに感じることができる環境です。
・平均年齢38歳　男女比 6:4　落ち着いて仕事に取り組める雰囲気です。

【当社のコア技術・プラットフォーム】

- **bit-MAP®**：微生物を1細胞ずつシーケンスし、従来の手法では到達できない高精度なゲノム解析を実現します。
- **bit-GEM**：bit-MAPを通じて構築された、30億以上の微生物遺伝子データベース。その多くは既存の公開データベースに存在しない未知のものです。ここから新規遺伝子や有用酵素を発掘し、ゲノムと表現型（フェノタイプ）を紐づけるモデルを学習させています。
- **bit-EVO**：大規模かつ高拡張性を持つDBTL（Design-Build-Test-Learn）サイクルを回すことで、ウェットラボで得られた各実験結果が、即座に次の設計精度を研ぎ澄ます。

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。

■受賞歴
2024年3月　Global Hypergrowth Tokyo プログラム2024
2023年11月 NEDO ディープテック・スタートアップ支援事業採択
2023年4月　XTC Japan 2023 優勝
2023年3月　NEDO グリーンイノベーション基金採択
2022年9月　大学発ベンチャー表彰2022　経済産業大臣賞
2021年8月　NEDO TRY採択

【雇用形態】
正社員
※試用期間あり、6ヶ月（条件の変更なし）

【給与】
想定年収　600～1200万円

（内訳）
月給：50～100万円
基本給部分：36.8～73.7万円+みなし残業時間45時間・時間外手当相当 13.1～26.2万円
※ご経験や前職給与等を考慮のうえ決定します

■人事評価：年2回

【就業時間】
基本9：00～18：00（実働8時間）専門業務型裁量労働制：1日の労働時間 8時間とみなす
休憩60分

残業：あり

【勤務地】
（雇入れ直後）東京都新宿区早稲田鶴巻町513 早稲田大学121号館
■本社/東京メトロ東西線：早稲田駅から徒歩2分

※勤務地変更の範囲：従業員の自宅
※※入社後、業務に慣れるまで（3ヶ月程度）はオフィス出勤、その後はチームと相談しリモート可。週1-3日オフィス勤務できる方が理想。

【休日休暇】

- 年間休日128日（今期 2026/1～2026/12 ※暦により変動有）
- 完全週休2日制
- 夏季休暇
- 年末年始
- 慶弔休暇
- 有給休暇（※入社半年後に10日を付与）

【待遇・福利厚生】

- 社会保険完備
- 交通費支給：50,000円/月 上限
- 博士取得サポート制度（博士後期課程進学時の学費支援）
- 社内勉強会・学会参加支援
- 社内イベント（年数回）

スキル・資格

【必須要件】

- 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
- バイリンガルチームで業務を遂行できるレベルの日本語または英語力（ビジネスレベル以上）

【歓迎要件】

- 微生物または原核生物のゲノミクスに関する実務経験
- ディープラーニング向けのデータ準備、およびモデルの学習・推論実行の経験（例：PyTorch）
- ゲノムまたはタンパク質の言語モデル（例：ESM、ProGen、Evo、GenomeOcean、Bacformer）およびそれらの入力データ形式に関する知識・理解
- ラボチームと連携し、LIMS（ラボ情報管理システム）においてサンプルやアッセイデータの登録・追跡を行った経験
- ワークフロー管理ツール（Nextflow、Snakemake、Airflow、Prefect、Dagsterなど）の使用経験

【選考について】

- 募集人数：1名
- 選考フロー：書類選考 >> 一次面接（HR） >> 二次面接（Head of Innovation・Principal AI Scientist ）>>リファレンスチェック>>最終面接（CEO/CSO/CTO）>> 内定可否

※最終面接以外の面接は、オンラインで実施予定です。
※状況により、面接・面談の回数が増える場合があります。
※他社選考によるスケジュール調整もご相談に応じます。

会社説明