

背景と目的

ホタルイカの「身投げ」とは、産卵期にあたる春（主に2月から5月）、富山湾の海岸にホタルイカが大量に押し寄せる現象である。この現象は幻想的な光景を生み出し、ホタルイカ掬いという独自の文化として地域に根付いており、シーズン中は県内外から多くの人々が訪れる。

現在、ホタルイカ掬いの愛好家が利用する情報源としては、個人の経験則や既存の掲示板サイトが主である。しかし、客観的なデータに基づく高精度な予測システムは確立されていない。

本研究の目的は、ホタルイカの身投げ量を定量的に予測する実用的なシステムを構築することである。

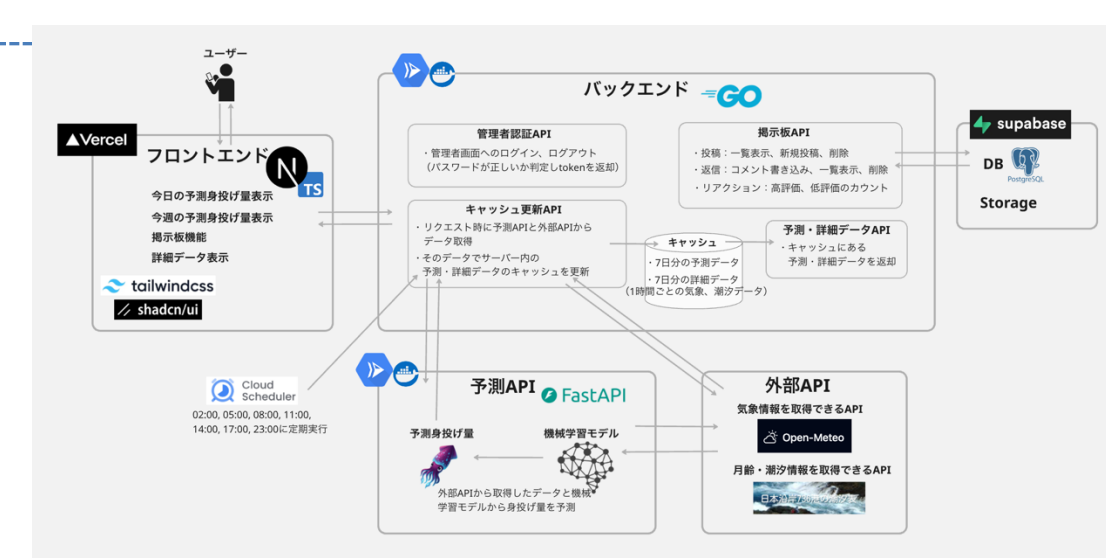
具体的には、以下の2点を達成することを目的とする。

- 過去の身投げ実績データと、それに対応する気象・海象（潮汐、月齢）データを収集・統合し、機械学習モデル（LightGBM）を用いて身投げ量を予測するモデルを構築する。
- 構築した予測モデルを組み込んだAPI、および予測情報と関連データ（天気、潮位等）、さらにユーザー間の情報交換を可能にする掲示板機能を提供するWebアプリケーションを開発し、ホタルイカ掬いを行う人々の利便性を向上させる。

研究方法

本研究は、「データセット構築」「モデル訓練」「予測API開発」「Webサイト構築」の4フェーズで実施した。

- データセット構築:** 過去のホタルイカ身投げ実績、気象データ（気温、降水量、風など）、潮汐、月齢データを収集し、機械学習用の訓練データセットを構築した。
- 機械学習モデルの訓練:** 構築したデータセットを用い、LightGBMによる予測モデルを訓練した。その際、月齢や日付の「周期性」をsin/cos変換で表現し、過去の気象情報を「ラグ特徴量」として追加する特徴量エンジニアリングを行った。モデルの検証は、データの時間的順序を維持する時系列交差検証を用い、ハイパーパラメータチューニングを実施した。
- 予測APIの開発:** 学習済みモデルをFastAPIでAPI化し、Cloud Runにデプロイした。このAPIは、外部の気象・潮汐APIから最新の予報データを取得し、訓練時と同一の前処理を適用して未来の身投げ量を予測する。
- Webサイトの構築:** Next.js（フロントエンド）とGo（バックエンド）で予測情報を表示するWebサイトを構築した。Goバックエンドが予測APIや外部APIの結果をメモリにキャッシュし、Cloud Schedulerによって3時間ごとに定時更新するアーキテクチャを採用した。データベースにはSupabase (PostgreSQL) を使用した。



結果と考察

予測モデルの評価

構築したLightGBMモデルの精度を評価するため、全データセットの最後の20%をテストデータとして使用し、予測値と実測値を比較した。

評価指標:

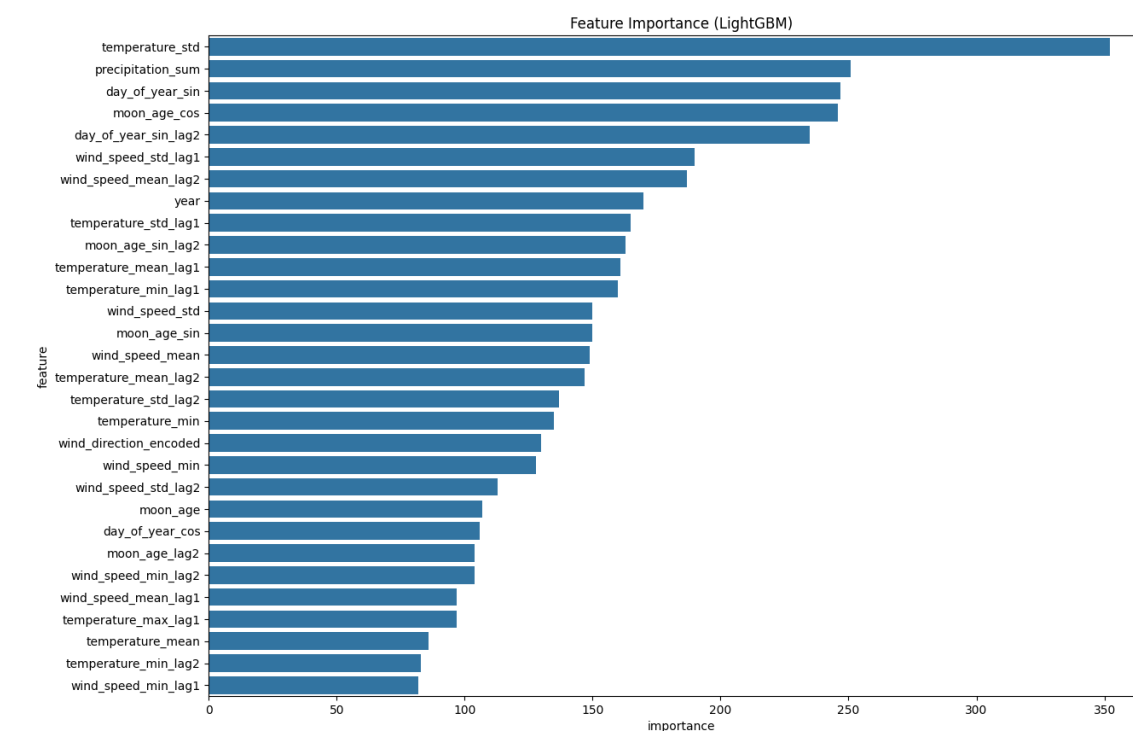
決定係数 (R^2): 0.7008

平均絶対誤差 (MAE): 0.0721

二乗平均平方根誤差 (RMSE): 0.1013

予測結果の可視化: 予測値（赤色）と実測値（青色）を時系列で比較したグラフ（記事の「予測結果と実測値の比較」図を参照）を作成した。

特徴量の重要度: 予測への寄与度が高い特徴量として、temperature_std（気温の標準偏差）、precipitation_sum（降水量）、day_of_year_sin（日付の周期性）、moon_age_cos（月齢の周期性）、wind_speed_std_lag1（1日前の風速の標準偏差）などが上位となった。



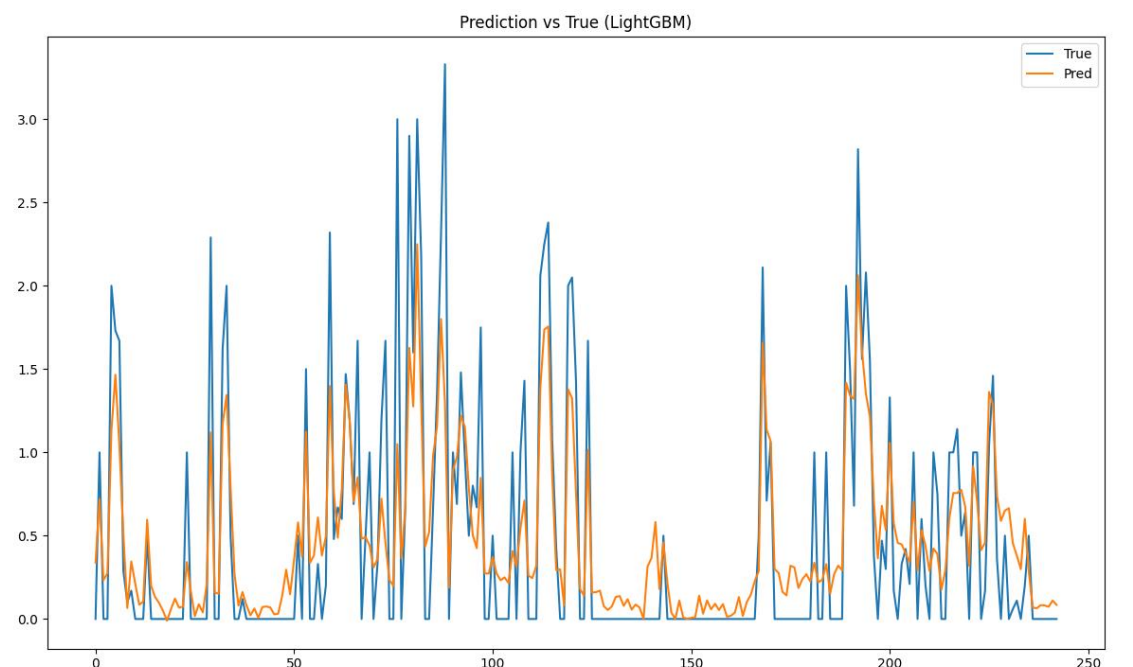
考察と課題

構築したLightGBMモデルのテストデータに対する決定係数 (R^2) は 0.7008 となり、一定の予測精度を達成した。特徴量の重要度分析からは、季節性（day_of_year）、月齢、気温、降水量といった、従来経験的に重要とされてきた要因が予測に寄与していることを定量的に確認できた。

一方で、モデルには以下の課題も明らかになった。

予測の精度: 全体の増減傾向は捉えられているものの、突発的な「爆湧き」のようなピーク値の予測精度には改善の余地がある。

外部データへの依存: モデルの予測精度は、入力データとして利用する外部の気象予報API自体の精度に直接依存する。気象予報が不正確な場合、本モデルの予測精度も低下する。



結論

本研究では、ホタルイカの身投げ量を予測するため、過去の時系列データに基づきLightGBMを用いた機械学習モデルを構築した。特徴量エンジニアリング（sin/cos変換、ラグ特徴量）を適用し、時系列交差検証によるチューニングを行った結果、テストデータにおいて決定係数 $R^2 = 0.7008$ という一定の予測精度を達成した。

さらに、この予測モデルを組み込んだFastAPIによる予測APIと、Next.js（フロントエンド）およびGo（バックエンド）を用いたWebアプリケーションを開発した。Goバックエンドでのキャッシュ機構やCloud Schedulerによる定時更新を導入することで、外部APIへの負荷を抑えつつ高速なレスポンスを実現する実用的なシステムを構築した。

今後の予定

- 予測精度の向上:** 海岸ごとの位置情報を特徴量として加え、より空間解像度の高い局所的な予測モデルを構築する。また、より高精度な気象・海象データを導入する。
- Webサイトの機能拡充:** オフシーズンでもユーザーが訪れる価値のあるコンテンツ（例：過去の身投げデータの統計・可視化ページなど）を追加し、サイトの継続的な利用を促進する。

⑧参考文献（任意）

⑧参考文献（任意） の内容を記載してください。