

TYDZIEŃ 3 - SPRAWOZDANIE

1. Dokładny opis projektu:

Potencjał biznesowy: realizowany przez nas projekt (zastosowanie metod ML w predykcji akcji Amazon.com) miałby ogromny potencjał biznesowy, ponieważ Amazon jest jednym z największych i najbardziej dynamicznych graczy na rynku e-commerce. Oto kilka aspektów tego projektu i jego potencjalnych korzyści:

- **Zarządzanie portfelem:** Prognozy mogą posłużyć do podejmowania decyzji dotyczących kupna, sprzedaży lub trzymania akcji Amazon.
- **Optymalizacja strategii inwestycyjnych:** Poprzez ciągłe doskonalenie strategii na podstawie wyników modelu ML, można osiągnąć lepsze wyniki inwestycyjne.
- **Personalizacja doradztwa finansowego:** Projekt może posłużyć do stworzenia systemu personalizowanego doradztwa finansowego opartego na predykcjach cen akcji Amazon. To może obejmować dostarczanie zindywidualizowanych rekomendacji inwestycyjnych i analizy portfela, które są dostosowane do preferencji, celów i poziomu tolerancji ryzyka każdego użytkownika.
- **Kształcenie inwestorów:** Poprzez udostępnianie wyników i wniosków modelu ML, projekt może także pomóc w kształceniu inwestorów na temat analizy rynku kapitałowego, funkcjonowania rynków finansowych i roli danych w podejmowaniu decyzji inwestycyjnych.

Potencjał naukowy:

- **Rozwój algorytmów ML:** Projekt taki stanowiłby pole do rozwoju zaawansowanych algorytmów uczenia maszynowego, które mogą efektywnie analizować i przetwarzać ogromne zbiory danych finansowych. Badania nad różnymi architekturami modeli ML, technikami uczenia się, funkcjami straty oraz metodami optymalizacji mogą przyczynić się do ulepszenia skuteczności predykcji cen akcji.
- **Integracja danych wielorodnych:** Aby skutecznie przewidywać ceny akcji Amazon.com, projekt ten wymagałby integracji danych z różnych źródeł, takich jak dane finansowe, dane społecznościowe, dane makroekonomiczne itp. Badania nad metodami integracji i wykorzystania wielorodnych danych mogą przynieść innowacyjne podejścia do analizy rynku.
- **Badania nad przyczynowością w analizie finansowej:** Analiza finansowa często koncentruje się na identyfikowaniu korelacji między różnymi zmiennymi. Jednak badania nad przyczynowością mogą pomóc w zrozumieniu, które czynniki rzeczywiście wpływają na zmienność cen akcji Amazon.com. Wykorzystanie metod uczenia maszynowego do badania przyczynowości w danych finansowych może przynieść nowe spojrzenie na rynki finansowe.

- **Badania nad strategiami inwestycyjnymi:** Projekt taki mógłby służyć jako platforma do badania różnych strategii inwestycyjnych i ich skuteczności w kontekście rynku Amazon.com. Analiza wyników strategii inwestycyjnych może dostarczyć cennych wskazówek dotyczących tego, jakie czynniki lub strategie są najbardziej obiecujące w kontekście inwestowania w akcje Amazon i rynki finansowe ogólnie.

Potencjalni partnerzy: firmy maklerskie, deweloperzy aplikacji giełdowych

Potencjalni odbiorcy: inwestorzy indywidualni, firmy inwestycyjne, instytucje finansowe, analitycy rynkowi

Metodologia działania:

1. **Zrozumienie problemu:** Pierwszym krokiem jest dokładne zrozumienie problemu i celów projektu. Należy zdefiniować, jakie dokładnie pytania chcemy zadać modelowi ML, jakie dane są potrzebne do tego celu oraz jakie metryki sukcesu zostaną użyte do oceny skuteczności modelu.
2. **Zbieranie danych:** Następnie należy zbadać dostępne źródła danych i zebrać odpowiednie dane finansowe, techniczne i inne, które mogą być istotne dla predykcji cen akcji Amazon.com. Ważne jest również zapewnienie jakości danych poprzez eliminację brakujących wartości, usuwanie błędnych danych i normalizację danych.
3. **Przygotowanie danych:** Dane muszą być odpowiednio przetworzone i przygotowane do analizy przez modele ML. Proces ten obejmuje skalowanie danych, normalizację i standaryzację danych oraz podział danych na zestawy treningowe, walidacyjne i testowe.
4. **Trening modeli:** Następnie należy przeprowadzić trening modeli na danych treningowych.
5. **Ocena modelu:** Po zakończeniu treningu modelu, należy ocenić jego skuteczność na danych walidacyjnych i testowych, używając odpowiednich metryk oceny.
6. **Optymalizacja i strojenie modeli:** Proces ten polega na optymalizacji parametrów modelu oraz jego hiperparametrów w celu uzyskania najlepszych wyników predykcji.
7. **Monitorowanie w czasie rzeczywistym:** Po otrzymaniu właściwie dostosowanych modeli konieczne jest monitorowanie ich skuteczności w odniesieniu do danych rzeczywistych.

Technologie, które zamierzamy wykorzystać w projekcie oraz do komunikacji:

- **Python**, w tym biblioteki Pandas, NumPy, Matplotlib, Seaborn, Scikit-Learn, TensorFlow, PyTorch, Keras; wybrane środowisko: **Jupyter Notebook**
- **ClickUp**: strona do organizacji pracy, podziału na zadania, określenia deadline'ów, podziału pracy oraz kontroli nad wykonywaniem działań
- **GitHub**: w celu kontroli wersji, kolaboracji, prowadzenia dokumentacji oraz organizacji plików
- **Google Docs**: w celu kolaboracji przy sporządzaniu sprawozdań
- **Messenger**: w celu komunikacji w czasie rzeczywistym

2. Koordynatorem projektu została Natalia Szczepkowska.

3. PLAN MIN&MAX:

Minimalny efekt, który będzie dla nas zadowalający, to:

- ☐ skompletowanie wszystkich zadań zgodnie z poniższym planem

▼ TO DO 15 ... + Add Task

Name

- ☐ 1.1 Określenie metodologii oraz rozplanowanie zadań
- ☐ 1.2 Zebranie oraz uzupełnienie danych
- ☐ 1.3 Data preprocessing, analiza zbioru danych
- ☐ 2.1 Zbudowanie modelu regresji liniowej
- ☐ 2.2 Zbudowanie modelu SVM
- ☐ 2.3 Zbudowanie modelu Random Forest Regression
- ☐ 2.4 Zbudowanie modelu Gradient Boosting Regressor
- ☐ 2.5 Zbudowanie modelu LSTM
- ☐ 2.6 Zbudowanie modelu GRU
- ☐ 2.7 Zbudowanie modelu CNN
- ☐ 2.8 Zbudowanie modelu ARIMA
- ☐ 3.1 Ocena modeli i wybór najlepszego z nich
- ☐ 3.2 Analiza skuteczności przewidywań w czasie rzeczywistym
- ☐ 4.1 Stworzenie prezentacji z opisem uzyskanych wyników, wniosków oraz oceną pracy zespołowej
- ☐ 4.2 Prezentacja

+ Add Task

- ☐ wyciągnięcie wniosków ze współczynników modelu
- ☐ wybranie najlepszego modelu

Maksymalny efekt to nie tylko realizacja tych zadań, ale także

- ☐ osiągnięcie wyjątkowej jakości pracy
- ☐ przekroczenie naszych oczekiwań

W kontekście czasowym:

Minimalne oczekiwania to:

- ☐ przestrzeganie ustalonych terminów
- ☐ spóźnienia nie większe niż 2 tygodnie

Maksymalne oczekiwania to:

- ☐ dokończenie pracy przed wyznaczonym terminem końcowym

WPŁYW EFEKTU:

MINIMALNEGO	NA MAKSYMALNY
Osiągnięcie tylko minimalnych wyników w jednym etapie	Ograniczenie zdolności w osiągnięciu maksymalnego
Spóźnienia	Doprowadzenie do wydłużenia czasu końcowego
Zmniejszenie motywacji zespołu	Wpływ na morale i osiąganie gorszych wyników w dalszej pracy
Niewystarczająco poświęcona uwaga szczegółom	Zwiększenie ryzyka błędu
Znikome twórcze myślenie	Ograniczenie innowacji, poprawy projektu