

TYDZIEŃ 2 - SPRAWOZDANIE

1. Skład grupy: Magdalena Lamczyk, Wiktoria Sarabon, Martyna Szawłowska, Jakub Galikowski, Natalia Szczepkowska
2. Temat projektu: Zastosowanie metod Machine Learning w przewidywaniu cen akcji firmy Amazon.com
[Ceny akcji Amazon](#)
[Przychody Amazon](#)
3. Harmonogram prac z podziałem na zadania:

ETAP PROJEKTU	OPIS
1. PREPROCESSING	Zebranie, oczyszczenie i uzupełnienie danych, analiza zbioru danych, przedstawienie podstawowych wykresów, analiza zależności
2. ZBUDOWANIE MODELI	Zbudowanie następujących modeli: a) liniowa regresja b) SVM (Support Vector Machines) c) Random Forest Regression d) Gradient Boosting Regressor e) LSTM (Long Short-Term Memory) (przydatny do analizy historycznych danych o cenach akcji) f) GRU (Gated Recurrent Units) g) CNN (Convolutional Neural Networks) h) ARIMA (Auto-Regressive Integrated Moving Average)
3. PORÓWNANIE I OCENA MODELI	Porównanie skuteczności i poprawności wyżej wymienionych modeli
4. DOPASOWANIE MODELI DO DANYCH PRZYSZŁYCH	Dopasowanie modelu do danych na następne miesiące oraz analiza ich poprawności w czasie rzeczywistym
5. PREZENTACJA WYNIKÓW	Prezentacja wyników predykcji cen akcji oraz ocena pracy w zespole

Jakie istnieją metody minimalizacji ryzyka projektu informatycznego? [Źródło](#)

- Metody eksperckie takie jak ankiety eksperckie, technika burzy mózgów, analiza SWOT, technika oceny i przeglądu programu (PERT), symulacje Monte Carlo itd., służą do identyfikacji, oceny i zarządzania ryzykiem. Na przykład, analiza SWOT może pomóc w identyfikacji potencjalnych zagrożeń (ryzyk) i szans (możliwości

zmniejszenia ryzyka), podczas gdy symulacje Monte Carlo mogą pomóc w ocenie wpływu różnych scenariuszy ryzyka.

- Metody analizy sieciowej takie jak technika CPM, PERT, GERT, są używane do planowania i kontroli projektu, co może pomóc w minimalizacji ryzyka poprzez identyfikację ścieżek krytycznych i potencjalnych opóźnień.
- Metoda czasowo – kosztowa CPM-COST, jest używana do zarządzania kosztami projektu, co może pomóc w minimalizacji ryzyka finansowego.

Ryzyka, które mogą wystąpić na różnych etapach projektu:

- Analiza Wymagań:
 - Ryzyko: Niedokładna lub niekompletna analiza wymagań może prowadzić do błędów w późniejszych etapach.
 - Minimalizacja: Regularne spotkania, konsekwentne monitorowanie zmian w wymaganiach, stosowanie technik prototypowania.
- Zbieranie Danych:
 - Ryzyko: Może wystąpić problem z dostępem do odpowiednich danych historycznych dotyczących cen akcji Amazon.com. Może to być spowodowane ograniczeniami praw autorskich, prywatności lub dostępności.
 - Minimalizacja: Wyszukanie alternatywnych źródeł danych historycznych, które mogą być dostępne publicznie lub w formie płatnych subskrypcji. Rozważenie skorzystania z wielu źródeł danych w celu zminimalizowania ryzyka związanego z zależnością od jednego dostawcy.
- Przygotowanie Danych:
 - Ryzyko: Oczyszczenie danych może być trudne, a niedokładne dane mogą wpłynąć na jakość modelu.
 - Minimalizacja: Ustalanie klarownych kryteriów dotyczących jakości danych, regularna weryfikacja danych, wykorzystanie narzędzi do oczyszczania danych.
- Projektowanie Modelu:
 - Ryzyko: Wybór niewłaściwego algorytmu lub nieodpowiednich cech może skutkować niską skutecznością modelu.
 - Minimalizacja: Testowanie różnych modeli, stosowanie technik doboru cech, regularna walidacja krzyżowa.
- Optymalizacja Modelu:
 - Ryzyko: Niemożność optymalizacji modelu może prowadzić do niskiej wydajności predykcyjnej.
 - Minimalizacja: Systematyczne strojenie parametrów modelu, analiza wpływu zmian na wyniki, wykorzystanie technik optymalizacji, takich jak Random Search czy Grid Search.
- Interpretacja Wyników:
 - Ryzyko: Wyniki mogą być trudne do interpretacji lub mogą nie przynieść oczekiwanych rezultatów.
 - Minimalizacja: Jasne zdefiniowanie celów projektu, dokładna dokumentacja procesu analizy, wizualizacja danych i wyników.
- Implementacja Systemu:
 - Ryzyko: Problemy podczas integracji modelu z systemem mogą prowadzić do awarii lub błędów.

- Minimalizacja: Testy integracyjne, stopniowa implementacja, współpraca między zespołem programistycznym a zespołem ds. danych.
- Testy:
 - Ryzyko: Niedostateczne testy mogą nie wykryć wszystkich błędów w systemie.
 - Minimalizacja: Planowanie testów systemowych, wykorzystanie testów jednostkowych i testów automatycznych, przeprowadzanie testów wydajnościowych.
- Szkolenie:
 - Ryzyko: Niedostateczne szkolenie użytkowników może prowadzić do błędnego korzystania z systemu.
 - Minimalizacja: Przygotowanie czytelnej dokumentacji, organizowanie szkoleń, tworzenie interaktywnych materiałów edukacyjnych.
- Wdrożenie:
 - Ryzyko: Problemy wdrożeniowe mogą prowadzić do opóźnień lub awarii systemu.
 - Minimalizacja: Stopniowe wdrażanie, planowanie awaryjnych procedur, testy przedwdrożeniowe.
- Monitorowanie i Utrzymanie:
 - Ryzyko: Brak ścisłego monitorowania może prowadzić do utraty skuteczności modelu z czasem.
 - Minimalizacja: Regularne aktualizacje modelu, monitorowanie jakości danych wejściowych, stosowanie automatycznych alertów.