**PLAN PRACY**

**Proponowany temat: Metody uczenia maszynowego usprawniające procesy detekcji oszustw z wykorzystaniem kart kredytowych.**

Wstęp

1. Charakterystyka Data Science

1.1. Wprowadzenie do konceptu nauki o danych

1.2. Business Intelligence(BI) – Analityka Biznesowa

1.3. Analiza porównawcza koncepcji Data Sciene oraz Business Intelligence(BI)

1.4. Eksploracyjna analiza danych

1.5. Idea(pojęcie?) uczenia maszynowego

2. Algorytmy oraz narzędzia wykorzystane w nadzorowanym uczeniu maszynowym

2.1. Klasyfikacja metod nadzorowanego uczenia maszynowego

2.2. Analiza implementacji przykładowych metod - charakterystyka

2.3. Przykłady zastosowań algorytmów uczenia maszynowego

2.4. Środowisko Azure Machine Learning Studio oraz inne wykorzystane narzędzia

2.5. Konkurencyjne środowiska oraz narzędzia dostępne na rynku

3. Projekt oraz implementacja modelu predykcyjnego wykrywającego transakcje noszące znamiona oszustwa

3.1. Czym jest detekcja oszustw transakcyjnych w kontekście kart kredytowych (Fraud Detection)

3.2. Charakterystyka użytego zestawu danych (dataset) – eksploracyjna analiza danych

3.3. Pomocniczy model optymalizacyjny wyłaniający najlepiej rokujących kandydatów spośród algorytmów oraz ich złożeń

3.4. Ocena jakości oraz możliwości dopasowania istniejących przykładowych implementacji metod..

3.5. Budowa modelu predykcyjnego.......................................................................................................

3.6. Porównanie istniejących metod z przeprowadzoną w ramach badań implementacją....................

3.7. Interpretacja uzyskanych wyników

3.8. Podsumowanie zagadnienia

Literatura

Lista rysunków

Lista tabel....................................................................................................................................................

Słowa kluczowe:

Uczenie maszynowe, oszutwa, Fraud detection, DSR?, detekcja anomalii, nadzorowane uczenie maszynowe, data science, cyberbezpieczeństwo, python, dataset, karta kredytowa, transakcje, numpy, pandas, azure ml studio, eda, BI, analiza, analityka biznesowa, big data, data mining, model predykcyjny, klasyfikator, regresja liniowa, SVM, R, R-studio, Jupyter notebook, Anaconda, google colab, IEEE, inzynieria cech, siec neuronowa, konwolucyjna siec neuronowa, gradient, drzewo decyzyjne, boosting, cross-validation, dataframe, perceptron, Bayes, sztuczna inteligencja, automatyzacja, optymalizacja, algorytmika, modelowanie, prototypowanie, propagacja wsteczna (backpropagation), estymator, entropia, preprocessing, random forest, statystyka.