

Cel zadania:

Przeanalizować działanie istniejącego modelu, ocenić jego skuteczność i zaproponować sposoby ulepszenia, wykorzystując ChatGPT jako narzędzie wspomagające analizę danych, inżynierię cech i projektowanie eksperymentów.

Część 1 – Analiza wstępna

1. Wczytaj dostarczony plik `biomaterials_dataset.csv` oraz model `svm_biomaterials_model.pkl`.
2. Oceń skuteczność modelu na danych testowych (dokładność, precision, recall, F1-score).
3. Wygeneruj **macierz pomyłek** i zidentyfikuj klasy, które są najczęściej mylone.
4. Zapytaj ChatGPT:
 - O interpretację najważniejszych cech w kontekście biomateriałów.
 - O możliwe przyczyny trudności modelu w rozróżnianiu wybranych klas.

Część 2 – Rozbudowa modelu

1. Przy pomocy ChatGPT opracuj listę dodatkowych cech (rzeczywistych lub syntetycznych), które mogą poprawić jakość klasyfikacji.
(np. gęstość materiału, przewodność cieplna, frakcja HA, stosunek modułu Younga do porowatości)
2. Dodaj przynajmniej 2 nowe cechy do danych (mogą być sztucznie wygenerowane w Pythonie według zaproponowanej zależności).
3. Wytrenuj nowy model SVM z tymi dodatkowymi cechami.
4. Porównaj wyniki nowego modelu z oryginalnym.

Część 3 – Eksperymenty i wnioski

1. Zapytaj ChatGPT o:
 - Alternatywne algorytmy, które mogą lepiej sprawdzić się w tym zadaniu.
 - Pomysły na walidację krzyżową i tuning hiperparametrów.
2. Zastosuj przynajmniej jedną z rekomendacji ChatGPT i oceń wpływ na wyniki.
3. Przygotuj **raport** zawierający:
 - Porównanie wyników modeli (oryginalny vs. ulepszony).
 - Wnioski z analizy błędów.
 - Rekomendacje dla zespołu badawczego (jakie dane warto pozyskać w przyszłości).

💡 **Wariant dla ambitnych:**

Zaproponuj ChatGPT, aby wygenerował **syntetyczny zbiór danych** z nowymi, realistycznymi cechami. Następnie porównaj jakość modeli trenowanych na danych oryginalnych vs. wzbogaconych.