# 📌 Analiza teoretyczna zadania — Multi-label classification

## 🔹 1. Definicja problemu

**Multi-label classification** (MLC) to problem uczenia maszynowego, w którym:

- Dla jednego przykładu $x$ chcemy przewidzieć **więcej niż jedną etykietę**.
- Każda etykieta $Y_k$ przyjmuje wartość binarną: $Y_k \in \{0, 1\}$.
- Przykład: dla obrazu przyrody $x$ etykiety mogą być: "mountain", "trees", "sea", itp.

**Porównanie:**

- **Binary classification**: przewidujemy jedną etykietę binarną.
- **Multi-class classification**: przewidujemy jedną etykietę wieloklasową spośród K klas.
- **Multi-label classification**: przewidujemy jednocześnie K etykiet binarnych.

## 🔹 2. Metody uczenia

### 📌 (a) Binary Relevance (BR)

- **Idea**: Dla każdego $Y_k$ trenujemy niezależny klasyfikator:

$$P(Y_k = 1|X)$$

- Każdy model ignoruje zależności między etykietami.
- Proste, ale nie uwzględnia współzależności między etykietami.

---

### 📌 (b) Classifier Chain (CC)

- **Idea**: Modeluje zależności między etykietami w sekwencji:
  - Model 1: $Y_1 \leftarrow X$
  - Model 2: $Y_2 \leftarrow X, Y_1$
  - Model 3: $Y_3 \leftarrow X, Y_1, Y_2$
  - ...
  - Model K: $Y_K \leftarrow X, Y_1, \ldots, Y_{K-1}$
- Każdy model wykorzystuje predykcje wcześniejszych modeli jako dodatkowe cechy.

**Uwaga**: Kolejność etykiet w łańcuchu ma wpływ na wyniki — niektóre etykiety mogą być lepszymi predyktorami innych.

---

### 📌 (c) Ensemble of Classifier Chains (ECC)

- **Idea**: Trenujemy kilka CC z różnymi permutacjami etykiet i uśredniamy ich predykcje.
- Zmniejsza wariancję i wrażliwość na kolejność etykiet.

## 🔹 3. Metryki oceny jakości

### 📌 (a) Subset Accuracy

- Czy wszystkie etykiety w danej próbce są poprawnie przewidziane:

$$\text{Subset Accuracy} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}\{y_i = \hat{y}_i\}$$

- Bardzo restrykcyjna metryka — wymaga pełnej zgodności.

---

### 📌 (b) Hamming Score

- Średnia dokładność etykiety:

$$\text{Hamming Score} = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{K}\sum_{k=1}^{K}\mathbf{1}\{y_{ik} = \hat{y}_{ik}\}\right)$$

- Bardziej tolerancyjna metryka — uwzględnia częściowo poprawne przewidywania.

---

## 🔹 4. Dane: emotions (OpenML)

- Zawiera:
  - $X$ — cechy numeryczne opisujące utwory muzyczne.
  - $Y$ — 6 binarnych etykiet emocji.
- Dane wczytywane przy użyciu:

```python
X, Y = fetch_openml("emotions", version=4, return_X_y=True)
```

- Etykiety binarne przekształcamy z "TRUE"/"FALSE" na 1/0.

## 🔹 5. Implementacja

### 📌 Binary Relevance (BR)

- 6 niezależnych klasyfikatorów, np. Logistic Regression.

### 📌 Classifier Chain (CC)

- Jeden łańcuch klasyfikatorów:
  - model 1 → model 2 (na wejściu ma też predykcje modelu 1) → model 3 → itd.

### 📌 ECC

- 5 permutacji CC, uśrednianie predykcji:

$$\hat{Y} = \text{średnia z } n \text{ CC} \geq 0.5$$

---