

Aufgabe

Klassifikation unbekannter Getränke für automatisiertes Produkt-Matching:

Ein Getränkeunternehmen verkauft über mehrere Plattformen hinweg und erhält Produktdaten von Drittanbietern oder Distributoren. Die Daten sind nicht standardisiert, sondern nur durch symbolische Eigenschaften beschrieben:

- "Bernsteinfarben, süß, spritzig"
- "Klar, bitter, spritzig"

Problem

Wie kann das Unternehmen automatisiert herausfinden, ob es sich bei diesen Beschreibungen um ein Bier, einen Wein oder eine Limonade handelt?

Dies ist wichtig für:

- Produktkategorisierung im Online-Shop
- Bestandsabgleich
- Personalisierte Empfehlungen
- Automatisierte Produkt-Labeling in Lieferketten

Installation – PyTorch (CPU)

In Python venv:

```
pip install torch --index-url https://download.pytorch.org/whl/cpu
```

Vorgehensweise

In der letzten Zelle des Jupiter-Notizbuchs wurde eines der Testgetränke als Beispiel genommen - „wahrscheinlich Limonade“ (siehe Screenshot). Jedes Attribut wurde geändert und die Klassifizierungs- und Ähnlichkeitswerte wurden berechnet. Erstellen Sie eine neue Zelle im Notizbuch und wiederholen Sie dies mit einem anderen Testgetränk.

```
# Geänderte Attributes für: 'wahrscheinlich Limonade'
test_drinks = [
    {"Farbe": "grün", "Geschmack": "fruchtig", "Alkoholgehalt": "alkoholfrei", "Temperatur": "warm"}, # Temperatur geändert
    {"Farbe": "grün", "Geschmack": "fruchtig", "Alkoholgehalt": "niedrig", "Temperatur": "kalt"}, # Alkoholgehalt geändert
    {"Farbe": "grün", "Geschmack": "herb", "Alkoholgehalt": "alkoholfrei", "Temperatur": "kalt"}, # Geschmack geändert
    {"Farbe": "klar", "Geschmack": "fruchtig", "Alkoholgehalt": "alkoholfrei", "Temperatur": "kalt"} # Farbe geändert
]

for drink in test_drinks:
    predicted, sims = classify(drink)
    print(f"Getestetes Getränk: {drink}")
    print(f"Vorhergesagte Kategorie: {predicted}")
    for cat, sim in sims.items():
        print(f"    {cat}: {sim:.4f}")
    print()
```

Vorgehensweise

1. Wie ermöglicht das Wissen, welcher Attributvektor zu einer besseren Klassifizierung beiträgt, die Erklärbarkeit? → Antworte Dies im Presentation.
2. Erstellen Sie 2 neue Prototypen von Getränken (z. B. Kaffee, Energydrinks usw.) und fügen Sie sie zusammen mit den vordefinierten in das Modell ein (bearbeiten Sie die im Screenshot gezeigte Teil). Wie behandelt das Modell die Klassifizierung mit den neu hinzugefügten Prototypen? Haben sich die Punktzahlen insgesamt erhöht oder verringert?

```
# Beispiel für jede Kategorie (kann beliebig erweitert werden)
prototypes_data = {
    "Wein": [("Farbe", "rot"), ("Geschmack", "herb"), ("Alkoholgehalt", "hoch"), ("Temperatur", "warm")],
    "Bier": [("Farbe", "braun"), ("Geschmack", "bitter"), ("Alkoholgehalt", "niedrig"), ("Temperatur", "kalt")],
    "Limonade": [("Farbe", "klar"), ("Geschmack", "süß"), ("Alkoholgehalt", "alkoholfrei"), ("Temperatur", "kalt")]
}
```