

# Machine Learning zum Cocktailmixen

Universität Stuttgart

Florian Grassl  
Fachsemester: 5

Betreuer: Moritz Fuchsloch, Kim Glumann  
Institut für Strahlwerkzeuge (IFSW)

Stuttgart, 24. Januar 2025

# Aufgabenblatt

Hier das offizielle Aufgabenblatt einfügen oder Text formulieren.

# Inhaltsverzeichnis

|                                      |          |
|--------------------------------------|----------|
| <b>Aufgabenblatt</b>                 | <b>I</b> |
| <b>1 Einleitung</b>                  | <b>1</b> |
| 1.1 Motivation . . . . .             | 1        |
| <b>2 Hauptteil</b>                   | <b>2</b> |
| 2.1 Methodik . . . . .               | 2        |
| 2.1.1 Datenvorverarbeitung . . . . . | 2        |
| 2.1.2 Modellierung . . . . .         | 2        |
| <b>3 Schluss</b>                     | <b>4</b> |
| 3.1 Zusammenfassung . . . . .        | 4        |
| <b>Literaturverzeichnis</b>          | <b>4</b> |

# Kapitel 1

## Einleitung

### 1.1 Motivation

Als OpenAI im November 2022 die erste Version des GPT-4-Modells veröffentlichte, stieg das Interesse der Öffentlichkeit an künstlicher Intelligenz und maschinellem Lernen stark an. Um die Komplexität und die Möglichkeiten von Machine Learning zu verdeutlichen, soll in dieser Arbeit ein Modell entwickelt werden, das das Mischen von Cocktails automatisiert. Dabei werden große Datensätze an Geschmacksprofilen kategorisiert und in fünf Geschmacksrichtungen (süß, sauer, bitter, fruchtig, würzig) aufgeteilt. Dies bietet die Möglichkeit, verschiedene Lernmodelle zu erlernen, zu testen und zu vergleichen.

# Kapitel 2

## Hauptteil

### 2.1 Methodik

#### 2.1.1 Datenvorverarbeitung

Zu Beginn wurde eine **JSON-Datei**-Datenbank angelegt mit den Zutaten und ihren Geschmacksprofilen. Diese Datenbank wurde in ein Python-Dictionary umgewandelt und in ein Dataframe überführt. Anschließend wurden die Daten normalisiert und in Trainings- und Testdaten aufgeteilt.

#### 2.1.2 Modellierung

Die Entwicklung des Systems zur Generierung personalisierter Cocktailrezepte basiert auf einer Kombination aus **Machine-Learning-Techniken** und **mathematischer Optimierung**. Der Prozess gliedert sich in mehrere Schritte:

#### Datenbasis und Nutzerprofil

Zunächst werden Zutaten mit ihren Geschmacksprofilen (süß, sauer, bitter, fruchtig, würzig) sowie Fragen zur Ermittlung der Nutzerpräferenzen aus einer **JSON-Datei** geladen. Dies dient dazu, ein individuelles, gewichtetes **Nutzerprofil** zu erstellen. Sowohl die Fragen, als auch die Zutaten lassen sich flexibel erweitern, um das System an verschiedene Anwendungsfälle anzupassen, oder ein noch genaueres Geschmacksprofil zu entwickeln. Der Nutzer durchläuft einen **Fragebogen**, der individuelle Geschmacksvorlieben erfasst. Jede Antwort wird in ein fünfdimensionales Vektorprofil übersetzt:

$$\mathbf{p}_{\text{user}} = \frac{1}{N} \sum_{i=1}^N \mathbf{q}_i,$$

wobei  $\mathbf{q}_i$  das Profil der  $i$ -ten Frage und  $N$  die Anzahl der beantworteten Fragen ist. Zusätzlich wählt der Nutzer aus, ob der Cocktail alkoholisch sein soll, und legt die maximale Zutatenanzahl  $k$  fest.

#### Autoencoder zur Profilrekonstruktion

Kern des Systems ist ein **Autoencoder**, der das Nutzerprofil komprimiert und rekonstruiert, um versteckte Muster in den Geschmacksdaten zu identifizieren. Die Wahl fiel

hier auf die Autoencoder-Variante, da diese als robuste, nicht-lineare und unüberwachte Methode die besten Ergebnisse liefert und noch zusätzlich als Feature-Extraktor für ein Supervised Learning dienen könnte. Die Architektur umfasst:

- **Encoder:** Reduziert das 5D-Profil auf einen 3D-Latent-Space:

$$\mathbf{z} = \text{Encoder}(\mathbf{p}_{\text{user}}),$$

- **Decoder:** Rekonstruiert das Profil aus dem Latent-Space:

$$\mathbf{p}_{\text{rekon}} = \text{Decoder}(\mathbf{z}).$$

Das Training erfolgt unüberwacht auf den Zutatenprofilen mit dem Loss:

$$\mathcal{L} = \|\mathbf{p}_{\text{rekon}} - \mathbf{p}_{\text{user}}\|^2.$$

## Optimale Zutatenmischung

Basierend auf  $\mathbf{p}_{\text{rekon}}$  wird die ideale Mischung durch Lösung des **Least-Squares-Problems** berechnet:

$$\min_{\mathbf{w}} \|\mathbf{A}\mathbf{w} - \mathbf{p}_{\text{rekon}}\|^2, \quad (2.1)$$

wobei:

- $\mathbf{A} \in \mathbb{R}^{5 \times M}$ : Matrix der Zutatenprofile (5 Geschmacksdimensionen,  $M$  Zutaten),
- $\mathbf{w} \in \mathbb{R}^M$ : Gewichtsvektor der Zutaten (Anteile).

Negative Gewichte werden auf null gesetzt, und nur die Top- $k$ -Zutaten (mit den höchsten Gewichten) bleiben erhalten. Die Gewichte werden aufsummiert und in Milliliter umgerechnet (Gesamtvolumen: 200 ml):

$$\mathbf{w}_{\text{final}} = \frac{\mathbf{w}_{\text{top-}k}}{\sum \mathbf{w}_{\text{top-}k}} \times 200.$$

# Kapitel 3

## Schluss

### 3.1 Zusammenfassung

In dieser Arbeit wurde ein System entwickelt, das mithilfe von Machine Learning und mathematischer Optimierung personalisierte Cocktailrezepte generiert. Der Ansatz kombiniert die Analyse von Geschmacksprofilen mit einem Autoencoder und einer Least-Squares-Optimierung, um individuelle Vorschläge zu erstellen. Die Ergebnisse zeigen, dass das System flexibel und skalierbar ist und sich für zukünftige Erweiterungen eignet.

# Selbstständigkeitserklärung

Ich versichere, dass ich diese Arbeit selbstständig verfasst habe.

Stuttgart, 24. Januar 2025

(Florian Grassl)