**DIPLOMARBEIT**

DermaAI: Intelligente Hautanalyse

Gesamtprojekt

Entwicklung eines KI-Modells zur Detektion und Klassifikation pigmentierten Hautläsionen und dessen Integration in eine Mobile App

Diplomarbeitsnummer  
5AHINF-24/25-DA11

**Datenbank, Verwaltung der medizinischen Daten**

Jonas Maier 5AHINF Betreuer: Dipl.-Ing. Dr.

Gerhard Gaube

**Mobile Integration, Frontend**

Jonas Bogensberger 5AHINF Betreuer: Dipl.-Ing. Dr.

Michael Wimmer

**KI-Modell in Python, API und Appanbindung**

Daniel Jessner 5AHINF Betreuer: Dipl.-Ing. Dr.

Gerhard Gaube

Schuljahr 2024/25

Abgabevermerk:

Datum: TT.MM.JJJJ übernommen von:

Inhalt

1. **Allgemeines / Projektübersicht**
   1. **Projektbeschreibung**
   2. **Projekt und Schnittstellen**
2. **Funktionale Anforderungen**
   1. **Use Cases**
3. **Nicht Funktionale Anforderungen**
4. **Projektplanung**
   1. **Variantenbildung**
   2. **Machbarkeitsstudie**
      1. **Projektziele**
      2. **Benötigte Ressourcen**
      3. **Entwicklungsmethodik**
      4. **Kommunikations- und Berichterstattungsstrategie**
   3. **Softwarearchitektur**
      1. **Aktivitätsdiagramme**
      2. **Sequenzdiagramme**
      3. **Komponentendiagramme**
      4. **Verteilungsdiagramme**
      5. **Softwarediagramme / Komponenten**
      6. **SW Programme**
      7. **SW Komponenten**
5. **Projektdurchführung**
   1. **Webscraping & labeln der Daten**
      * 1. **Datenquelle und Beschaffung**

**Da die Quelle der Daten eine besonders wichtige Rolle im Zusammenhang mit dem Trainieren der KI ist, war es besonders wichtig auf zuverlässliche Daten vertrauen zu können. Da die Auswahl an medizinisch anerkannten Medizinischen Datenbanken nicht gerade gering ist, habe ich weiter recherchiert und mich letztendlich für das** [**Isic-Archive**](https://www.isic-archive.com) **entschieden. Überzeugt hat mich die generell hohe Qualität der Bilder und die Ähnlichkeit der Aufnahmeart. Die Datenbeschaffung habe ich mittels Python-Webscraping realisiert, jedoch dazu weiter im nächsten Kapitel.**

* + - 1. **Datenbereinigung und Vorverarbeitung**

**Zum Webscrapen, damit wir die Bilddaten für unser Projekt verwenden können, habe ich mithilfe von verschiedenen Python Modulen gearbeitet. Der erste Schritt war die Links und die Diagnose mittels RegEx herauszufiltern und mittels Request an den Link des Bildes letztendlich das Bild selbst zu erhalten. Im nächsten Schritt war es vonnöten diese Bilder zuzuschneiden und für die KI vorzubereiten. Einfach gesagt habe ich mittels einer OpenCV-Funktion die Anomalien in einem Bild ausgeschnitten und zu Graustufen verändert. Das führte letztendlich zu inkonsistenten Bildern in Graustufen, welche jedoch aufgrund von fehlender Skalierung nicht für das Machine-Learning verwendbar sind. Eine weiter wichtige Information für diesen Abschnitt ist auch, dass ich für Performancezwecke auch die Scripte aufgeteilt habe und deshalb auf einen „shared Memory“, auch geteilten Speicher, zugreifen musste. Dieser geteilte Speicher ermöglichte es mir einfach große Datenmengen, wie die Bilddaten selbst, zwischen Scripten auszuteilen.**

* + - 1. **Labeln der Daten**

**Das Labeln der Daten habe ich mittels JSON-Objekts umgesetzt, was heißt, dass das es ein Attribut Picture gibt, welches logischerweise auch ein Bild als Numpyarray repräsentiert. Das zweite Attribut der JSON ist die Diagnose, welches die Diagnose des Arztes selbst ist. Wichtig hinzuzufügen ist auch, dass ich einen großen Teil meiner Zeit in das Zuschneiden der Bilder auf das nötigste gelegt habe.**

* + 1. **Datenbank-Backend**
       1. **Datenmodell und Struktur**

**Als Datenbank haben wir eine No-SQL Datenbank verwendet, da es für unser Vorhaben perfekt in unser Bild gepasst hat. Durch die Funktionsweise sind wir nicht an ein reguläres Schema gebunden, sondern können einfach mittels JSON-Objekte an die Datenbank senden. Das Backend meiner API verwendet jedoch doch ein Schema, welches ein Bild und die Diagnose in einem JSON-Dokument haben will. Mit dieser Konfiguration ist es uns möglich eine einheitliche persistente Speicherung der Daten zu erzielen, was uns später vieles vereinfachen wird. Eine Implementation von Usern ist nicht Teil unserer Muss-Ziele, jedoch wollen wir, wenn genug Zeit vorhanden ist ein solches System wie in unseren Soll-Zielen umsetzten und somit weitere Funktionalität zu unserer Applikation bringen.**

* + - 1. **Technologien und Tools**

**Um uns das bestmögliche Backend zu machen habe ich auf modernste Technologien wie das Typescript-Framework** [**Adonisjs**](https://adonisjs.com) **gesetzt. Um eine Verbindung zur Datenbank aufzubauen habe ich auf das Package Mongoose gesetzt, welches mich nun bereits durch das 3 Projekt im Schulischen Umfeld begleitet. Der Techstack von Adonisjs ist nicht gering, bietet jedoch für sehr viele Situationen wichtige Funktionalität, welche das Entwickeln um einiges einfacher machen. Mithilfe der gut formulierten und instandgehaltenen Dokumentation ist das Einleben in das Framework besonders leicht gemacht.**

* + - 1. **Datenbankoperationen**
      2. **Verschlüsselung der gespeicherten Daten**

**#**

* + - 1. **Backup und Disaster-Recovery-Pläne**

**#**

* + - 1. **Test und Validierung**

**#**

* + 1. **Vorbereitung**
    2. **Aktionsphase**
    3. **Ende**

1. **Installation / Software deployment**
2. **Projektabschluss**

# Literaturverzeichnis

# Abbildungsverzeichnis