



**FOM Hochschule für Oekonomie & Management**

Hochschulzentrum Münster

## **Hausarbeit**

im Studiengang Big Data & Business Analytics

zur Erlangung des Grades eines

**Master of Science (M. Sc.)**

über das Thema

**Ansätze zu Optimierung und Automatisierung des OSMI-Index**

von

**Fiete Ostkamp, Verena Rakers und Artur Gergert**

Betreuer : Prof. Dr. Rüdiger Buchkremer

Matrikelnummer : 557851, 536491, 562394

Abgabedatum : 13. August 2022

# Inhaltsverzeichnis

<b>Abbildungsverzeichnis</b>	<b>III</b>
<b>Tabellenverzeichnis</b>	<b>IV</b>
<b>Abkürzungsverzeichnis</b>	<b>V</b>
<b>Symbolverzeichnis</b>	<b>VI</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Problemstellung . . . . .	1
1.2 Zielsetzung . . . . .	1
1.3 Vorgehen . . . . .	2
<b>2 Theoretische Grundlagen</b>	<b>3</b>
2.1 Online Sensory Marketing Index . . . . .	3
2.2 Deep Learning . . . . .	4
2.2.1 Image Classification . . . . .	4
2.2.2 Convolutional Neural Network . . . . .	4
2.3 Natural Language Processing . . . . .	5
2.3.1 Named Entity Recognition . . . . .	6
2.3.2 Vorgehen bei der NER . . . . .	6
<b>3 Praktische Umsetzung</b>	<b>9</b>
3.1 OSMI-Index . . . . .	9
3.2 Datenvorbereitung . . . . .	10
3.3 Modellentwicklung . . . . .	11
3.3.1 NER-Modell . . . . .	11
3.3.2 Image Classification Modell . . . . .	11
3.4 Klassifizierung von Webseiten . . . . .	12
3.4.1 Datenerhebung . . . . .	12
3.4.2 Datenverarbeitung . . . . .	13
3.4.3 Training des Modells . . . . .	14
<b>4 Fazit</b>	<b>14</b>
<b>Anhang</b>	<b>15</b>

## Abbildungsverzeichnis

Abbildung 1: Aufbau des Convolutional Neural Networks . . . . .	5
Abbildung 2: Beispiel der Named Entity Recognition (NER) Technik . . . . .	7
Abbildung 3: Zielbild des Consulting-Projekts . . . . .	9
Abbildung 4: Annotationsstand Textlabeling . . . . .	11
Abbildung 5: Annotationsstand Textlabeling . . . . .	11
Abbildung 7: Vergleich der Trainings- und Validierungsmetriken . . . . .	14

## **Tabellenverzeichnis**

## Abkürzungsverzeichnis

<b>CNN</b>	Convolutional Neural Network
<b>FN</b>	False Negative
<b>FP</b>	False Positive
<b>FP</b>	False Positive
<b>NER</b>	Named Entity Recognition
<b>NLP</b>	Natural Language Processing
<b>OSMI</b>	Online Sensory Marketing Index
<b>TP</b>	True Positive

## **Symbolverzeichnis**

# 1 Einleitung

Für Unternehmen wird es immer wichtiger die potenziellen Kunden gezielt anzusprechen und zum Kauf anzuregen. Eine Möglichkeit kann hierbei das multisensorische Marketing bieten, welches zur Beeinflussung des Kaufverhaltens mehrere Sinne anspricht. In einer Umfrage der Mood Media Corporation<sup>1</sup>, die unter 10.000 Menschen weltweit durchgeführt wurde, hat ergeben, dass 75% der Kunden sich länger in einem Geschäft aufhalten würden, wenn durch Musik oder Düfte eine positive Atmosphäre geschaffen wird.<sup>2</sup> Multisensorisches Marketing anzuwenden kann demnach einen positiven Effekt auf das Kaufverhalten haben.

## 1.1 Problemstellung

Die Anwendung von multisensorischem Marketing findet zurzeit mehrheitlich in stationären Geschäften statt. Da der Online-Handel weiterhin an Wachstum gewinnt sollte auch hier geprüft werden inwiefern multisensorisches Marketing eingesetzt werden kann. Ein von K. Hamacher entwickelter Index, der Online Sensory Marketing Index Online Sensory Marketing Index setzt an dieser Thematik an. Es ist ein Evaluationsmodell, welches Webseiten anhand ihrer multisensuellen Kommunikationsqualitäten bewertet. Allerdings bietet der Online Sensory Marketing Index noch Raum für Optimierungen. So wird der Online Sensory Marketing Index bisher manuell und subjektiv ermittelt und die Ergebnisse lediglich in einer Tabelle dargestellt. Des Weiteren wird bei der Beurteilung einer Webseite nicht nach Produktwebseiten und nicht-Produktseiten unterschieden und die Textanalyse mittels TF-IDF analysiert die sensorischen Wörter nicht im Kontext des gesamten Satzes.

## 1.2 Zielsetzung

Aus der Problemstellung lässt sich daher nachfolgende Zielsetzung für diese Ausarbeitung ableiten. Es soll eine Webseite erstellt werden, in derer unterschiedliche Produktwebseiten automatisiert hinsichtlich ihrer textlichen Ansprache an die Sinne des Konsumenten aufgeführt werden. Durch eine Evaluierung wird dargelegt wie gut oder wie schlecht die Webseite die Sinne Tasten, Riechen, Hören, Schmecken und Sehen anspricht. Darauf aufbauend könnten dann die Webseiten-Betreiber die Webseiten optimieren, um die Kunden besser anzusprechen.

---

<sup>1</sup> steiner2022.

<sup>2</sup> moodmedia2019.

### **1.3 Vorgehen**

Ein erster Schritt in der Umsetzung der Zielsetzung ist es die relevanten produktspezifischen Webseiten herauszufiltern. Mittels einer Bildklassifizierung soll dies umgesetzt werden, um diese für die weitere Bearbeitung verwenden zu können. Andere Unterseiten zum Unternehmen oder zu Karrieremöglichkeiten fließen somit nicht mehr in eine Bewertung ein. Im zweiten Schritt werden durch Natural Language Processing (NLP) die Texte auf den Produktseiten analysiert. Es werden gezielt Wörter - unter Einbezug des Satzkontextes - markiert, die auf eines der fünf Sinne hindeuten. Eine Beurteilung dieses Ergebnisses wird anhand eines Benchmarkings ermittelt und auf einer Webseite zur Verfügung gestellt.



## 2 Theoretische Grundlagen

### Quellen überall ergänzen

Im folgenden Teil der Ausarbeitung werden die theoretischen Grundlagen zum Verständnis der Ausarbeitung erläutert. Zunächst wird der Online Sensory Marketing Index (OSMI) als Hauptaugenmerk erklärt. Es folgen der Oberbegriff Deep Learning und die dazugehörige Methode der Bildklassifizierung. Darauf folgt das NLP mit der NER.

### 2.1 Online Sensory Marketing Index

Der Fokus in dieser Ausarbeitung liegt auf dem Online Sensory Marketing Index. Es handelt sich dabei um ein Bewertungsinstrument, welches im Jahr 2018 entwickelt wurde.**Zitat** Ziel des OSMI ist es, dass E-Commerce Webseiten im Hinblick auf deren sensorische Gestaltung bewertet werden und dementsprechend Verbesserungspotentiale aufgezeigt werden. Unter einer sensorischen Gestaltung werden Parameter gefasst, die sich nach den Sinnen eines Menschen richten. Diese sind die Haptik, die Olfaktorik, die Akustik, die Gustatorik und die Visualität.**Zitat** Die Bewertung im OSMI wird anhand verschiedener Indikatoren der einzelnen Sinne vorgenommen. Nachfolgend sind beispielhafte Indikatoren der einzelnen Sinne aufgeführt, die im OSMI Berücksichtigung finden:

- Haptik: 3D-Bilder, Video, Endowment-Effekt, ...
- Olfaktorik: Mentale Simulation von Düfte, Schlüsselbegriffe für Düfte
- Auditive: Töne/Musik, Sprecher/Stimme, Schlüsselbegriffe für Akustik, ...
- Gustatorische: Mental Simulation des Schmeckens, Verwendung von Farbschemata, Produktoption & Oberflächen, ...
- Visuelle: Farbgebung, dynamische Bilder, Oberfläche, ...

Die Evaluation der einzelnen Indikatoren wird in Form einer Tabelle dargestellt. Ein Beispiel ist am Ende dieses Kapitels aufgeführt **ref label der tabelle**. Für eine Webseite wird jeder Indikator der fünf Sinne einzeln bewertet und auf Basis dessen letztendlich für jeden der fünf Parameter ein Wert zwischen 0 und 1 ermittelt. Eine Gewichtung zwischen den einzelnen Indikatoren wird dabei nicht vorgenommen.

Im letzten Schritt zur Ermittlung des OSMI wird über die Parameter der fünf Sinne das arithmetische Mittel berechnet. Dieser Wert liegt ebenfalls zwischen 0 und 1 und bildet den OSMI einer Webseite. Dabei gilt, dass die Webseite erfolgreicher die Sensorik anspricht, wenn der OSMI-Index näher an einer 1 liegt. Liegt der Wert näher an einer 0

werden schlussfolgernd wichtige Elemente zur Ansprache der Sensorik nicht genutzt. Die nachfolgende Tabelle zeigt ein Beispiel aus der Masterarbeit von K.Hamacher zum OSMI der Webseite [www.mymuesli.de](http://www.mymuesli.de).

mymuesli Beispiel ergänzen? Hier noch was zu schreiben?

## 2.2 Deep Learning

Die Image Classification ist ein Anwendungsfall für die Technik des Deep Learning. Als Teilgebiet des Machine Learning bedient sich das Deep Learning mehrschichtiger neuronaler Netze, welche auf Basis großer Datenmengen bereits Erlerntes mit neuen Inhalten verknüpft und somit wiederum erneut lernt. **Zitat** Deep Learning macht es möglich, dass eine Maschine selbstständig die Gewichtungen seines verwendeten Modells verbessert. Anwendungsbereiche für Deep Learning können beispielsweise die Gesichts-, Objekt- oder Spracherkennung sein. **Zitat je Anwendungsfall** Die Anwendung der Objekterkennung wird nachfolgend beschrieben.

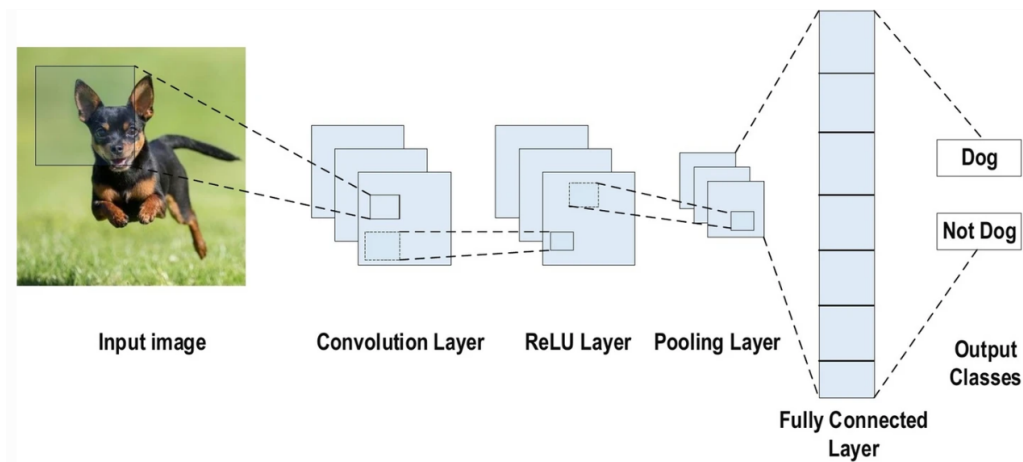
### 2.2.1 Image Classification

Bei der Image Classification handelt es sich um eine Technik, bei der ein Objekt auf einem Bild einer bestimmten Klasse zugeordnet wird. Als Ziel kann die genaue Identifizierung der Bildmerkmale verstanden werden. Es gibt verschiedene Methoden, die zur Image Classification eingesetzt werden können wie die Supervised Classification, die Unsupervised Classification oder auch die Convolutional Neural Networks (CNNs). **Zitat** Letzteres wird am häufigsten für den Anwendungsfall herangezogen. **Zitat**

### 2.2.2 Convolutional Neural Network

Das CNN ist eine besondere Form eines künstlichen neuronalen Netzes. Es zeichnet sich durch mehrere Schichten wie den Convolutional- Layer, den Pooling Layer und den Fully-Connected Layer aus. Die Architektur ist in der Abbildung reffig:CNN dargestellt.

**Abbildung 1: Aufbau des Convolutional Neural Networks**



Quelle: alzubaidiReviewDeepLearning2021

In den Schichten der Convolution werden die Eingabebilder analysiert und Merkmale (zum Beispiel Linien oder Kanten) der Bilder erkannt und extrahiert. Die Verarbeitung dieser Merkmale erfolgt in Matrizen, die Feature Maps genannt werden.

Die Feature Maps als Ausgabe der Convolutional Layer werden darauf in den Pooling Layer eingegeben. Die Größe der Bilder wird in dieser Schicht reduziert und durch Methoden wie dem Max-Pooling nur die wesentlichen Merkmale weitergegeben. Alle anderen Informationen sind für die Verarbeitung überflüssig und werden verworfen. Ergebnis dieser Schicht ist die gleiche Anzahl an Feature Maps, allerdings in komprimierter Form.

Der Fully-Connected Layer bildet den Abschluss der CNN-Architektur. Jeder Knoten in der Ausgabeschicht ist hier direkt mit einem Knoten der vorgelagerten Schicht verbunden. Mit Hilfe einer Linearkombination und einer Aktivierungsfunktion wird schlussendlich das Bild klassifiziert. Die Zuordnung zu einer Klasse erfolgt dabei auf Basis einer Wahrscheinlichkeit.

## 2.3 Natural Language Processing

Die Named Entity Recognition ist eine Technik, welche dem Oberbegriff des Natural Language Processing zugeordnet wird. Das NLP beschäftigt sich mit der Verarbeitung und dem Verständnis der menschlichen Sprache durch einen Computer. **Zitat** Ziel dieser Techniken ist es eine direkte Kommunikation zwischen dem Menschen und dem Computer mit Hilfe der Sprache herzustellen. Die Fachgebiete der Computerlinguistik, Informatik, Kognitionswissenschaft und künstlicher Intelligenz werden hier vereint. Anwendungsfälle können beispielsweise die Extraktion der Bedeutung von Sätzen oder Satzteilen oder auch die

Erkennung von Satzzusammenhängen sein. Neben Techniken wie der Sentiment-Analyse und der Spracherkennung gehört auch die Named Entity Recognition zu den Funktionen des NLP.

### 2.3.1 Named Entity Recognition

Die Named Entity Recognition hat zur Aufgabe automatisiert Entitäten in einem Textdokument zu finden und zu klassifizieren.**Zitat** Letzteres geschieht anhand zuvor definierter Kategorien. Eine Entität muss dabei nicht unbedingt nur aus einem einzelnen Wort bestehen, sondern kann auch eine Reihe von Wörtern umfassen, solange sie sich auf dieselbe Sache beziehen.

### 2.3.2 Vorgehen bei der NER

In der NER-Technik werden zwei Schritte durchgeführt, um zu Ergebnissen zu gelangen:

1. Erkennen einer Entität
2. Kategorisierung der Entität

Im ersten Schritt wird geprüft, ob eine Wortfolge eine Entität bildet. Die Anfangs- und Endgrenzen der Entitäten werden hierbei festgelegt. Der zweite Schritt hat zum Ziel die zuvor definierte Entität in eine der zuvor definierten Klassen einzuordnen. Häufig verwendete Klassen sind beispielsweise Orte, Namen oder Organisationen, welche unter die generischen Kategorien fallen.**Zitat** Des Weiteren gibt es domänenspezifische Kategorien wie zum Beispiel Proteine, Enzyme oder Gene.

Zur Durchführung der beiden Schritte gibt es verschiedene Ansätze, die angewendet werden können. Ausgehend von den annotierten Datensätzen, die beispielsweise aufgrund von manuell erstellten Regeln oder auf Basis von Kontextähnlichkeiten generiert werden, werden Modelle mithilfe von Machine Learning entwickelt. Für zuvor ungesehene Daten ermitteln die Modelle Vorhersagemodelle zur Erkennung und Kategorisierung der Entitäten. Eine weitere Möglichkeit ist der Einsatz von Deep Learning zur NER. In diesem Ansatz können auch nicht-lineare Zusammenhänge erkannt und gelernt werden.**Zitat** Ein Beispiel für ein Textdokument, in dem die NER Technik angewendet wurde, ist in Abbildung **ref xy** dargestellt.

## Abbildung 2: Beispiel der NER Technik

performance vehicle that excites with unique power delivery, tantalizing engine sound, excellent handling and a  
 •SOUND

direct driving sensation. the proverbial exclusivity also comes as standard, as do outstanding quality and

precise workmanship. in brief: all mercedes-amg models embody the levels of engineering and craftsmanship  
 •PRODUCT

performance vehicle that excites with unique power delivery, tantalizing engine sound, excellent handling and a  
 •SOUND

direct driving sensation. the proverbial exclusivity also comes as standard, as do outstanding quality and

precise workmanship. in brief: all mercedes-amg models embody the levels of engineering and craftsmanship  
 •PRODUCT

*Quelle: Eigene Darstellung*

Die NER-Technik kann in unterschiedlichen Bereichen angewendet werden. Zum Beispiel können durch die Anwendung die Antwortzeiten im Kundendienst verringert werden, indem die Anfragen zuvor kategorisiert oder direkt dem zuständigen Mitarbeiter zugeordnet werden. **Zitat**

Um die Qualität der NER-Technik zu bewerten, werden die Kennzahlen Precision, Recall und der F-Score hinzugezogen. Zur Ermittlung dieser muss zunächst die Anzahl der Entitäten mit verschiedenen Ausprägungen ermittelt werden (False Positive (FP), False Negative (FN), True Positive (TP)).

- FP: Eine Entität wurde erkannt, obwohl sie keine darstellt
- FN: Eine Entität wurde nicht erkannt, obwohl sie eine darstellt
- TP: Eine Entität wurde richtig erkannt

Die Precision zeigt das Verhältnis zwischen richtig erkannten Entitäten und der Gesamtheit der identifizierten Entitäten an. Die Formel lautet wie folgt:

$$Precision = \frac{TP}{(TP + FP)} \quad (1)$$

Der Recall stellt den Anteil der richtig erkannten Entitäten an der Gesamtheit aller möglichen Entitäten dar.

$$Recall = \frac{TP}{(TP + FN)} \quad (2)$$

Der F-Score ist die Kennzahl, der die Precision und den Recall zu einem harmonischen Mittel vereint:

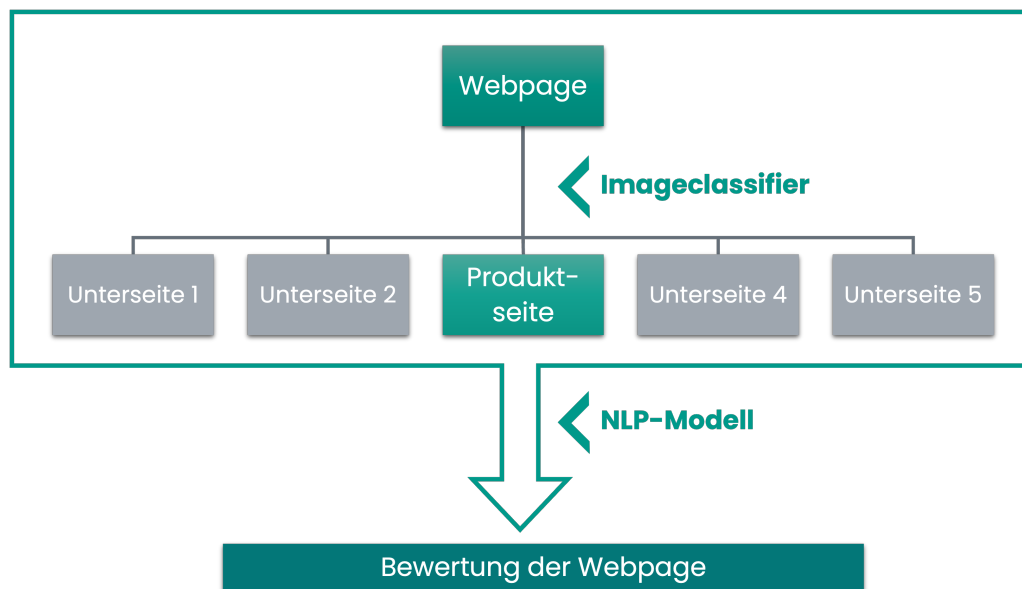
$$F - Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (3)$$

### 3 Praktische Umsetzung

In der praktischen Umsetzung werden die Grundlagen auf den Anwendungsfall angewendet. Hierfür wird zunächst der OSMI-Index erläutert. Es folgen die Bildklassifizierung und die Umsetzung der Named Entity Recognition (NER). Zum Abschluss des praktischen Teils folgt ein Kapitel mit Consultinganteil.

#### 3.1 OSMI-Index

**Abbildung 3: Zielbild des Consulting-Projekts**



Quelle: Eigene Darstellung

Um den OSMI-Index zu optimieren und zu automatisieren, wurde die in Abbildung 3 Pipeline aufgestellt mithilfe dessen eine Website in Bezug auf den OSMI verarbeitet wird. Der erste Schritt stellt dabei eine Vorverarbeitung der zu analysierenden Website dar, indem ein Image Classifier die Haupt- und Unterseiten der Website in Produkt- und Nicht-Produktseiten unterteilt. Dies ist aus Sicht der Autoren dieser Hausarbeit ein essenzieller Schritt, da so sichergestellt wird, dass die Bewertung des OSMI-Index lediglich auf Basis der relevanten URLs vorgenommen wird. Im darauf folgenden Schritt wird auf die textlichen Inhalte der relevanten Produktseiten ein NLP-Modell angewendet, welches die Entitäten des OSMI-Index erkennt und aus zählt. Im abschließenden Schritt werden die Ergebnisse der Website-Analyse im Rahmen eines Dashboards visualisiert. Nachfolgend wird die Umsetzung der soeben beschriebenen Schritte im Detail beschrieben und erläutert.

## 3.2 Datenvorbereitung

Sowohl in der Literatur als auch in bereits etablierten, gut entwickelten und veröffentlichten Modellen ist das Gebiet des Sensory Marketings sehr rar vertreten. Das bedeutet auch, dass im Hinblick auf den jüngst publizierten OSMI-Index ebenfalls bisher wenig technische Entwicklungen vorgenommen worden sind. Zwar existieren in der Literatur diverse NLP-Modelle wie beispielsweise das generelle Spacy-Modell oder der sogenannte BERT und von beiden erwähnten Modellen Spezialisierungen wie beispielsweise SciSpacy, BioBERT, ESG-BERT, etc; allerdings ist bisher kein NLP-Modell auf Marketingkontexte trainiert und publiziert worden, sodass die Notwendigkeit zur Entwicklung eines eigenen Modells im Kontext des Consulting-Auftrags bestand. Um im Vorfeld der Website-Bewertung jedoch lediglich nur die relevanten URLs mit Produktinhalt zu selektieren, wurde für den zu entwickelnden Image Classifier ebenfalls ein von Grund auf neu zu entwickelndes Modell benötigt, da es innerhalb der Literatur bislang kein derartig existierendes Modell publiziert wurde auf das zurückgegriffen werden konnte.

Die Trainingsdaten wurden mithilfe von Doccano generiert. Doccano ist ein open source Tool, welches zur Annotation von Text und Bilddaten entwickelt wurde. Mit diesem Instrument ist es also möglich Textpassagen und Bilder hochzuladen und Annotationen vorzunehmen, um Daten zur Entwicklung von Aufträgen wie Named-Entity-Recognition, Textzusammenfassungen, Sentimentanalysen, Image Classification, etc. zu generieren.

Im Rahmen des Consulting-Projekts wurden zwei Doccano-Projekte gestartet. Ein Projekt wurde für das Annotieren von Textpassagen zur Entwicklung eines Named-Entity-Recognition Modells erstellt. Im zweiten Projekt wurden Screenshots von Produktseiten und Nicht-Produktseiten annotiert.

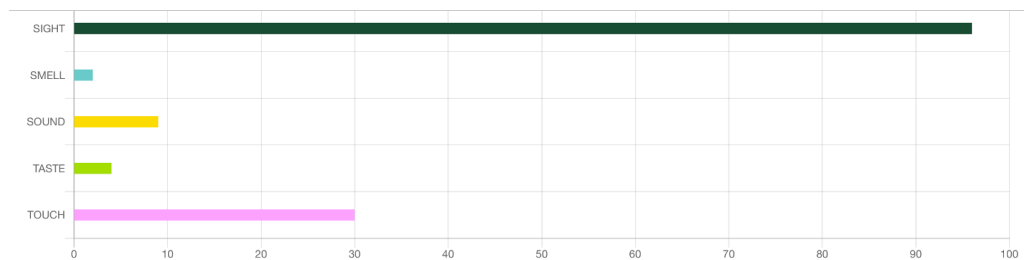
Die Textpassagen, die annotiert wurden, sind im Rahmen des Consulting-Projekts zur Verfügung gestellt worden und entstammen Webscraping-Ergebnissen von Vorgängerprojekten. Hier wurden die fünf Sinne

- Sight
- Smell
- Sound
- Taste
- Touch



als Label festgelegt und konnten im Rahmen des Annotationsprozesses ausgewählt werden. Die in Abbildung 4 dargestellten horizontalen Balken geben den Annotationsstand wieder, auf Basis dessen das Modell zur Erkennung von Entitäten entwickelt wurde. Hier ist zu erkennen, dass die Entität „Sight“ am häufigsten annotiert worden ist, während die Trainings- und Testdaten für die Entität „Smell“, „Sound“ sowie „Taste“ kaum vertreten sind.

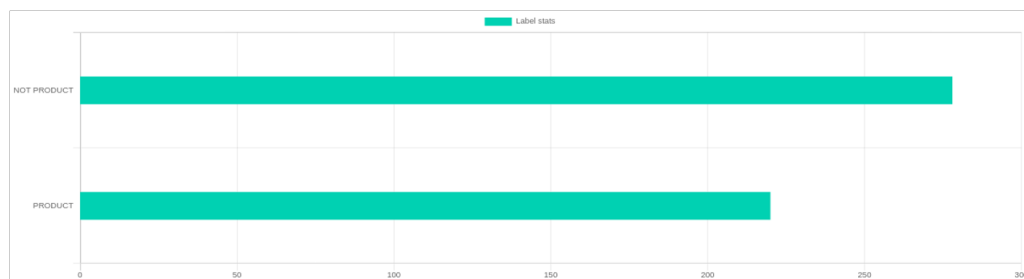
**Abbildung 4: Annotationsstand Textlabeling**



Quelle: Eigene Darstellung

Die zu annotierenden Bilddaten von Produkt- und Nicht-Produktseiten wurden im Rahmen des behandelten Projektes selbständig generiert. Abbildung 5 zeigt, dass zur Modellentwicklung für den Image Classifier eine fast Gleichverteilung zwischen Produkt- und Nicht-Produktseiten zugrunde gelegt wird. Der Gesamtdatensatz beläuft sich auf rund 500 annotierten Bildern von Homepages.

**Abbildung 5: Annotationsstand Textlabeling**



Quelle: Eigene Darstellung

### 3.3 Modellentwicklung

#### 3.3.1 NER-Modell

#### 3.3.2 Image Classification Modell

-Gewichtung der Indikatoren wird im OSMI nicht vorgenommen, weil nicht zweifelsfrei zu argumentieren welcher mehr oder weniger bewertet wird (bisher keine Forschungsarbeiten

dazu)

-Indikatoren werden einzeln bewertet & letztendlich fünf Parameter zw. 0 & 1 liegen vor. Daraus ergibt sich Gesamt-Index (ohne Gewichtung, arithmetischer Durchschnitt) für jede Webseite > auch zw. 0 & 1 und ist dann der OSMI-Index

-Je näher der OSMI an einer 1, desto erfolgreicher spricht die Webseite die Sensorik an. Wert nahe 0 erhält wichtige Elemente entsprechend nicht & erfüllt die Indikatoren nicht

-BSP eines OSMI-Indexes: Bild einfügen aus der Masterarbeit mymuesli.de oder so

möglicher Aufbau: 1. Bildklassifizierung a.Datenset b.Umsetzung 2. NER a. Datenset b. Umsetzung 3. Zusammenführung 4. Consultingteil

### **3.4 Klassifizierung von Webseiten**

Zur Differenzierung von Produkt- und nicht-Produkt Webseiten soll ein Klassifizierungsmodell erstellt werden. Im Folgenden wird das Vorgehen mit den Schritten der Datenerhebung (3.4.1), der Datenverarbeitung (3.4.2) und des Modelltrainings (3.4.3) näher beschrieben.

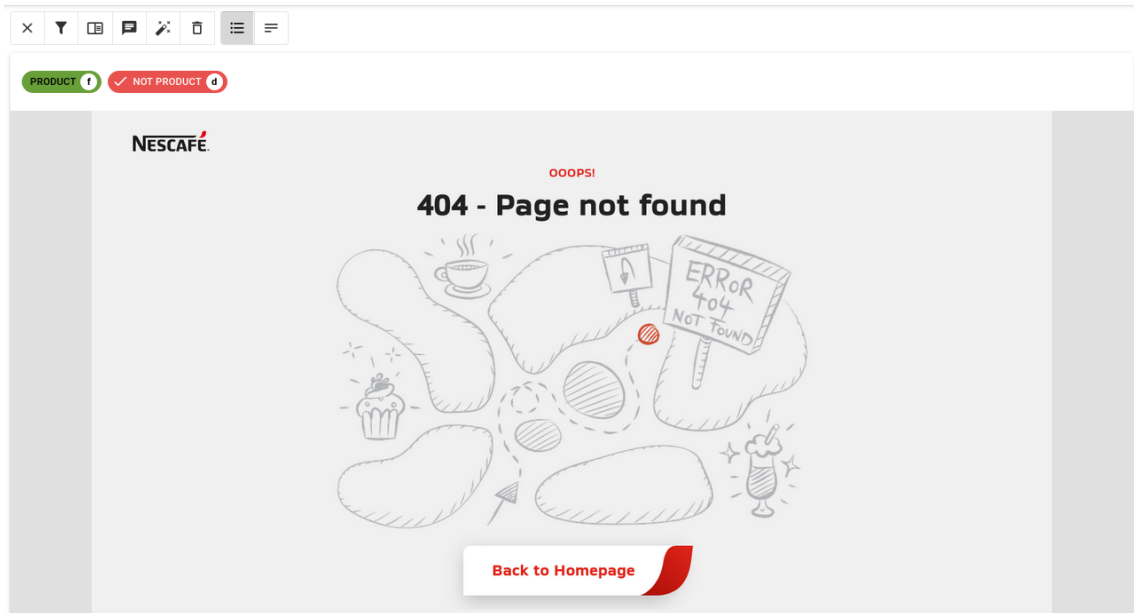
#### **3.4.1 Datenerhebung**

Aus vergangenen Arbeiten von Studierenden steht bereits eine Datengrundlage zur Verfügung mit Textauszügen, dem HTML Quelltext und heruntergeladenen Bildern von Unternehmenswebseiten. Für die Klassifizierung werden jedoch Bilder beziehungsweise Screenshots benötigt, die die gerenderte, visuelle Struktur der Webseite erfassen. Dies ist in den bestehenden Daten nicht gegeben und macht daher eine erneute Erhebung notwendig.

Für diesen Zweck wurde zunächst ein Webcrawler auf Basis von Scrapy<sup>3</sup> erstellt, der alle in der Vergangenheit gescrapten Webseiten aufruft und einen Screenshot der jeweiligen Seite erstellt. Auf diese Weise wurden circa 10.000 Bilder gesammelt, die als Eingangsdaten für das Modell verwendet werden können.

---

<sup>3</sup> zotero-328.

**Abbildung 6: Bildklassifizierung in Doccano**

Da es sich bei der Klassifizierung um eine Supervised Learning Methode handelt, müssen die Daten daraufhin den Kategorien *Product* und *Non Product* zugeordnet werden. Dafür werden die Daten in die Annotationssoftware Doccano<sup>4</sup> geladen und dann, wie in Abbildung 6 dargestellt, den Kategorien zugeordnet.

### 3.4.2 Datenverarbeitung

Im Schritt der Datenverarbeitung werden die Eingangsdaten in eine Form gebracht, in der sie für das Training des Modells verwendet werden können. Dazu werden die Bilder anhand des Exports der Annotationen aus Doccano in den Labeln gleichnamige Ordner verschoben.

Da die Screenshots als *.png* Bilder vorliegen, verfügen sie über einen Alpha-Channel der für die Verwendung im Modell hinderlich ist. Aus diesem Grund wurde dieser in einem weiteren Schritt entfernt. Darüber hinaus werden die Bilder im ersten Layer des Modells auf eine Größe von 224x224 Pixeln herunterskaliert, was eine gängige Größe für Convolutional Neural Networks darstellt.<sup>5</sup>

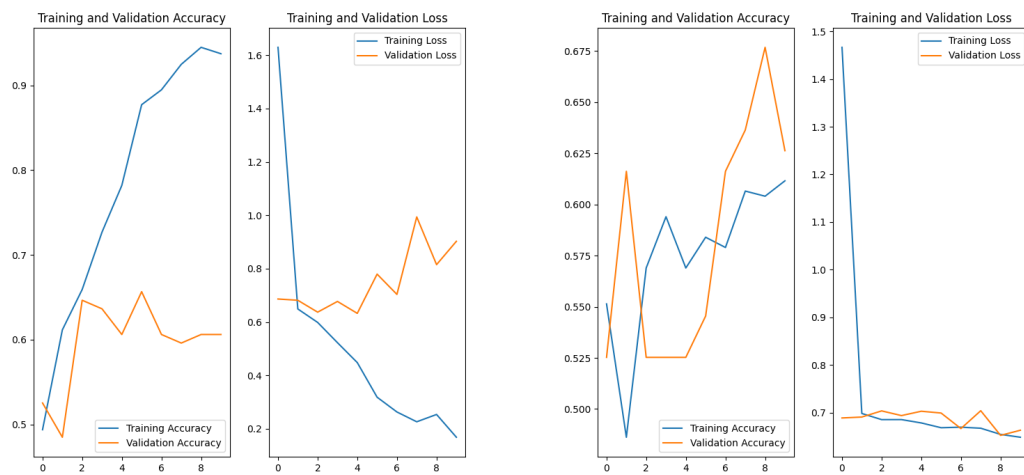
<sup>4</sup> doccano-2018.

<sup>5</sup> ghosh2019.

### 3.4.3 Training des Modells

Nachdem die Daten, wie in Kapitel 3.4.2 beschrieben, verarbeitet wurden, können Sie im weiteren Verlauf für das Training eines Modells verwendet werden. Dazu wurden die Daten wie im Deep Learning üblich, in Train und Validation Sets aufgeteilt um einen späteren Vergleich auf unbekannten Daten zu ermöglichen. Abbildung 7 stellt jeweils die Trainings-

**Abbildung 7: Vergleich der Trainings- und Validierungsmetriken**



**(a) Accuracy**

**(b) Loss**

*Quelle: Eigene Darstellung*

und Validierungsmetriken für die *Accuracy* und den *Loss* gegenüber. Es ist erkennbar, dass die Graphen in beiden Abbildungen relativ früh auseinander gehen, was darauf hindeutet, dass ... Insgesamt muss gesagt werden, dass die erzielten Genauigkeiten und Losses weit von einem idealen Ergebnis entfernt liegen.

## 4 Fazit

## Anhang

### Anhang 1: Beispielanhang

Dieser Abschnitt dient nur dazu zu demonstrieren, wie ein Anhang aufgebaut sein kann.

#### Anhang 1.1: Weitere Gliederungsebene

Auch eine zweite Gliederungsebene ist möglich.

### Anhang 2: Bilder

Auch mit Bildern. Diese tauchen nicht im Abbildungsverzeichnis auf.

#### Abbildung 8: Beispielbild

Name	Änderungsdatum	Typ	Größe
 abbildungen	29.08.2013 01:25	Dateiordner	
 kapitel	29.08.2013 00:55	Dateiordner	
 literatur	31.08.2013 18:17	Dateiordner	
 skripte	01.09.2013 00:10	Dateiordner	
 compile.bat	31.08.2013 20:11	Windows-Batchda...	1 KB
 thesis_main.tex	01.09.2013 00:25	LaTeX Document	5 KB

---

## Ehrenwörtliche Erklärung

Hiermit versichere ich, dass die vorliegende Arbeit von mir selbstständig und ohne unerlaubte Hilfe angefertigt worden ist, insbesondere dass ich alle Stellen, die wörtlich oder annähernd wörtlich aus Veröffentlichungen entnommen sind, durch Zitate als solche gekennzeichnet habe. Ich versichere auch, dass die von mir eingereichte schriftliche Version mit der digitalen Version übereinstimmt. Weiterhin erkläre ich, dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde/Prüfungsstelle vorgelegen hat. Ich erkläre mich damit **einverstanden/nicht einverstanden**, dass die Arbeit der Öffentlichkeit zugänglich gemacht wird. Ich erkläre mich damit einverstanden, dass die Digitalversion dieser Arbeit zwecks Plagiatsprüfung auf die Server externer Anbieter hochgeladen werden darf. Die Plagiatsprüfung stellt keine Zurverfügungstellung für die Öffentlichkeit dar.

Münster, 13.8.2022

(Ort, Datum)

A handwritten signature in black ink, consisting of a large, stylized 'H' followed by a series of loops and a final flourish.

(Eigenhändige Unterschrift)