

Bachelorprojekt

Sentimentanalyse für IMMANUEL

Timo Probst

Betreut durch Dr. Felix Lindner

Inhaltsverzeichnis

1. Einführung
2. Sentiment Detection
3. Vorverarbeitung
 - 3.1. Die Wort-Datenbank
 - 3.2. Stammformreduktion
4. PAD-Emotion Skala
5. Pipeline
 - 5.1. Ansatz 1
 - 5.1.1. Funktionsweise
 - 5.1.2. Vor- und Nachteile
 - 5.2. Ansatz mit Smoothing
 - 5.2.1. Funktionsweise
 - 5.2.2. Vor- und Nachteile
6. Evaluation
 - 6.1. F-Score
 - 6.2. Resultate
7. Diskussion
8. Zusammenfassung
9. Referenzen

1 Einführung

Das HERA-Projekt (Hybrid Ethical Reasoning Agents) ist ein von Prof. Martin Mose Bentzen (Technical University of Denmark) und Dr. Felix Lindner (Albert-Ludwigs-Universität Freiburg) vorangetriebenes Projekt, dessen Ziel es ist Werkzeuge zu entwickeln, mit denen Roboter in der Lage sind moralische Entscheidungen zu treffen und diese mithilfe ethischer Prinzipien wie dem Utilitarismus oder dem Prinzip der Doppelwirkung zu begründen (F. Lindner, M. M. Bentzen & B. Nebel, 2003). Im Rahmen des Projektes entstand IMMANUEL (Interactive Moral Machine based on multiple Ethical principles), ein anthropomorpher Roboter, welcher die Software des HERA-Projektes implementiert und benutzt wird um Mensch-Roboter-Interaktionen bei moralischen Diskussionen zu untersuchen (F. Lindner & M. M. Bentzen, 2017).



Abbildung 1: IMMANUEL ¹

Um mit Menschen kommunizieren zu können benutzt IMMANUEL die Sprachsynthesesoftware ‚MaryTTS‘², die eingegebenen Text in Audio umwandelt, welches dann dem Gesprächspartner vorgespielt wird. Bisher klang der von IMMANUEL ausgesprochene Text sehr monoton und emotionslos. Das lässt IMMANUEL zum einen roboterhaft erscheinen, was bei einem anthropomorphen Roboter nicht erwünscht ist. Zum anderen sorgt die unnatürlich klingende Stimme für ein anstrengenderes Zuhören.

Unser Ziel war es deshalb IMMANUELS Aussprache menschlicher zu machen, indem wir Informationen über das emotionale Befinden IMMANUELS in der Sprache simulieren. MaryTTS ermöglicht uns die Annotation des Textes mit den Emotionen Neutral, Glücklich, Traurig, Wütend und ‚Poker‘. Markiert man einen Textausschnitt mit einer Emotion, so wird die Aussprache des Textes entsprechend verändert. Da die Annotation nicht für jeden Text manuell erfolgen soll, brauchen wir einen Weg automatisch Emotionen zu detektieren und den Text damit ohne menschliches Zutun zu analysieren. Die Forschungsrichtung ‚Sentiment Detection‘ beschäftigt sich unter anderem mit der Untersuchung von Texten auf Emotionen.

¹ Quelle: Grafik von ‚The HERA Project‘, Web, Zugegriffen am 21.02.2018, in http://www.hera-project.com/wp-content/uploads/2016/12/robot2_web-750x275.jpg

² <http://mary.dfki.de/>

2 Sentiment Detection

Sentiment Detection ist die Wissenschaft, bei der man, unter Benutzung von Textanalyse, Untersuchung von natürlicher Sprache und Computerlinguistik, versucht Empfindungen in einem Medium, zum Beispiel einem Text, zu bestimmen.

Besonders das Web2.0, das Internet in dem jeder durch Posts, Nachrichten und Tweets seinen eigenen Content erstellen kann, hat das Interesse an der Sentiment Detection gestärkt. Für Firmen wie Facebook und Google ist es enorm relevant zu wissen wie User über Produkte und Formen posten, um diesen Usern dann optimal auf sie zugeschnittene Werbung präsentieren zu können.

Praktisch alle Methoden in der Sentiment Detection basieren auf einer Basismenge an bekannten Wörtern, die mit Emotionen oder Polaritäten bereits annotiert worden sind. Einfache Methoden wie der ‚Bag of Words‘-Ansatz bauen direkt auf dieser Wort-Datenbank auf, indem, um eine Schätzung für die Emotion zu erhalten, Wörter aus dem Text direkt mit Wörtern aus der Basismenge verglichen werden (J. Martineau & T. Finin, 2009). Wörter aus dem Text, für die eine annotierte Emotion gefunden werden kann, bestimmen die Emotion des Textes. Beispielsweise könnte bei dem Satz „Willkommen! Was für ein wundervoller Tag.“ ‚Willkommen‘ und ‚wundervoller‘ in der Basismenge mit der Emotion *fröhlich* annotiert worden sein, was dazu führt, dass der Satz ganze als *fröhlich* detektiert wird.

Fortgeschrittenere Methoden analysieren zusätzlich die Satzstruktur von zu untersuchenden Sätzen. Mithilfe von Regeln der Grammatik werden die Wörter im Satz in einem Baum organisiert, welcher dann die Satzstruktur beschreibt (T. Wilson, J. Wiebe & P. Hoffmann, 2005). Informationen aus einzelnen Wörtern werden dadurch in Kontext gesetzt. Vorteil dieses Ansatzes ist, dass kontextabhängige Bezüge zwischen Objekten erfasst werden können, was eine präzisere Emotionsanalyse ermöglicht. Das kann durch ein einfaches Beispiel verdeutlicht werden: Der Satz „Mir geht es nicht gut.“ würde bei einer einfachen ‚Bag of Words‘-Methode aufgrund des Wortes ‚gut‘ als positiv bezeichnet werden. Nur ein Ansatz, der die Satzstruktur mitbeachtet, kann erkennen, dass das Wort ‚nicht‘ die Polarität von ‚gut‘ umkehrt.

Diese Herangehensweise ist deutlich komplexer als eine einfache ‚Bag of words‘-Analyse, jedoch kann man so eine Präzision von 75-95% an richtig annotierten Sätzen erreichen (Tetsuya Nasukawa & Jeonghee Yi, 2003). Hierbei findet jedoch meistens nur eine Unterteilung in ‚positive Polarität‘ und ‚negative Polarität‘ eines Satzes statt, was die Komplexität des Problems verringert.

Während aktuell analytische Methoden, wie die Untersuchung von Satzstrukturen kombiniert mit einer großen Basismenge an Wörtern, noch bessere Resultate erzielen, besteht die Hoffnung, dass in Zukunft die künstliche Intelligenz ein präziseres Analysieren von Emotionen ermöglicht. Wir nutzen im Folgenden für die Analyse des Textes einen simplen ‚Bag of Words‘-Ansatz.

3 Vorverarbeitung

3.1 Die Wort-Datenbank

Wir benötigen also eine Datenbank, die deutsche Wörter auf Emotionen abbildet. Dazu haben wir zwei passende bereits vorhandene Datenbanken gefunden:

Die erste Datenbank ist ein Nebenprodukt einer Emotionsverlauf-Analyse von Roman Klinger, Surayya Samat Suliya und Nils Reiter (IMS Universität Stuttgart). Die Datenbank beinhaltet die insgesamt 4735 Wörter auf die sieben Emotionen Ekel, Freude, Furcht, Trauer, Überraschung, Verachtung und Wut abbildet.

Die zweite Datenbank ist eine englische Wort-Datenbank von Saif M. Mohammad und Peter D. Turney (Institute for Information Technology, National Research Council Canada.) die durch Crowdfunding entstanden ist. Diese Datenbank beinhaltet die acht Emotionen Wut, Erwartung, Ekel, Angst, Freude, Trauer, Überraschung und Vertrauen mit insgesamt 8265 Wörtern. Die einzelnen Wörter in der Datenbank wurden per Google-Übersetzer ins Deutsche übersetzt um dann mit der deutschen Datenbank zusammengefügt werden zu können.

3.2 Stemming

Bei einem Abgleich von Wörtern aus der Datenbank mit Wörtern aus einem zu untersuchenden Text stößt man auf ein Problem: Die Wörter ‚Wundervoller‘ aus „*Wundervoller Tag*“ und ‚wundervoll‘ aus der Datenbank sind nicht string-gleich miteinander. Das verhindert ein einfaches Matching von Wörtern aus dem Text mit Wörtern aus der Datenbank.

Eine naive Lösung wäre es eine Datenbank zu konstruieren, die alle möglichen Formen eines Wortes enthält. Eine solche Herangehensweise würde mit enormen Speicherverbrauch quittiert werden, da für jeden Wortstamm mehrere Wörter gespeichert werden müssen. Zur Veranschaulichung genügt es sich das Wort ‚wundervoll‘ herzunehmen und mögliche Modifikationen des Wortes zu betrachten: ‚wundervoller‘, ‚wundervolles‘, ‚wundervolle‘, ‚wundervollen‘, ‚wundervollem‘, ‚Wundervolles‘, ‚Wundervolle‘, ‚Wundervollen‘, ‚Wundervollem‘... Die Datenbank würde ein Vielfaches mehr an Speicherplatz verbrauchen, will man alle Abwandlungen eines Wortes mit abspeichern.

Zu einer besseren Lösung gelangt man durch Stemming (deutsch: Stammformreduktion). Hier wird jedes Wort durch sukzessive Anwendung von sprachspezifischen Reduzierungsregeln auf seinen Wortstamm zurückgeführt. So werden alle oben aufgelisteten Formen von ‚wundervoll‘ auf den Wortstamm ‚wundervoll‘ reduziert. Auch Groß- und Kleinschreibung wird beachtet, indem Anfangsbuchstaben zu Kleinbuchstaben gemappt werden.

Stemming erlaubt uns Worte aus unserer Datenbank mit Worten aus dem Text zu Matchen, sollten diese den gleichen Wortstamm besitzen. Die Stemming-Algorithmen funktionieren in den meisten Fällen sehr zuverlässig, selten kommt es zu dem Problem, dass zwei eigentlich unterschiedliche Worte auf den selben Stamm reduziert werden. Als Beispiel: ‚Bar‘ und ‚Bär‘

werden beide auf ‚bar‘ reduziert, sodass das Wort ‚Bar‘, welches eigentlich mit *Freude* annotiert wurde, auch mit *Angst* assoziiert wird. Solche Fehler beeinflussen die Stabilität unserer Emotionen-Analyse, da es sich hierbei jedoch um Ausnahmefälle handelt, nehmen wir diese eventuellen Uneindeutigkeiten in Kauf.

4 PAD-Emotion Skala

Wir sind jetzt in der Lage Wörter mit Emotionen aus unserer Datenbank zu verknüpfen. Da die so generierten Emotionen im Folgenden dargestellt, verglichen und verarbeitet werden müssen, benötigen wir eine Repräsentation der Emotionen, die handbarer und flexibel genug ist um das zu erlauben: Die PAD-Emotion Skala ist ein von Albert Mehrabian and James A. Russell entwickeltes psychologisches Modell um alle Emotionen durch einen 3-dimensionalen kontinuierlichen Vektorraum zu beschreiben. Die 3 Dimensionen des Raumes sind dabei *Freude* (engl.: pleasure), *Erregung* (engl.: arousal) und *Dominanz* (engl.: dominance). Jeder Emotion wird ein 3-dimensionaler Vektor mit Werten zugewiesen, welcher die Freude, Erregung und Dominanz der Emotion repräsentiert.

Wir haben die PAD-Werte für unsere Emotionen aus der Arbeit von S. Zhang et al. (2010) übernommen. Alle Werte kann man aus der Tabelle in Abbildung 2 entnehmen. Da das PAD-Modell kontinuierliche Werte zulässt, ist es möglich Emotionen zu kombinieren und dadurch neue Emotionen zu erhalten. Auch unterstützt MaryTTS die Angabe von Emotionen im PAD-Modell. Beides werden wir uns bei unseren späteren Emotions-Analysen zu Nutze machen.

| Emotion | Freude | Erregung | Dominanz |
|--------------|--------|----------|----------|
| Wut | -0.59 | 0.08 | 0.47 |
| Ekel | -0.59 | -0.01 | 0.4 |
| Angst | -0.08 | 0.18 | -0.39 |
| Fröhlichkeit | 0.63 | 0.40 | 0.29 |
| Neutral | 0.03 | -0.04 | -0.07 |
| Traurigkeit | -0.28 | -0.12 | -0.37 |
| Überraschung | 0.41 | 0.55 | 0.19 |

Abbildung 2: PAD-Werte

5 Pipeline

5.1 Ansatz 1

Funktionsweise

Unser erster Ansatz zur Annotierung des Textes ist recht unkompliziert: Der zu untersuchende Text wird in einzelne Sätze aufgeteilt. Jeder Satz wird unabhängig mit einer ‚Bag of Words‘-Analyse untersucht, wobei jedes Wort auf ein Vorkommen in der Emotions-Datenbank überprüft wird. Bei einem Treffer merken wir uns, dass sich ein Wort mit der entsprechenden Emotion in unserem Satz befindet. Nachdem alle Wörter im Satz untersucht wurden, wird die dominante Emotion, also die Emotion, die als häufigste im Satz vorkommt, als finale Emotion bestimmt. Da wir mit unserer Datenbank 9 verschiedene Emotionen detektieren können, MaryTTS jedoch nur die Emotionen Neutral, Glücklich, Trauer, Wütend und Poker zulässt, müssen wir jeder Emotion aus unserer Datenbank genau eine, für MaryTTS darstellbare, Emotion zuordnen. Die Zuordnung kann man aus Abbildung 3 entnehmen.

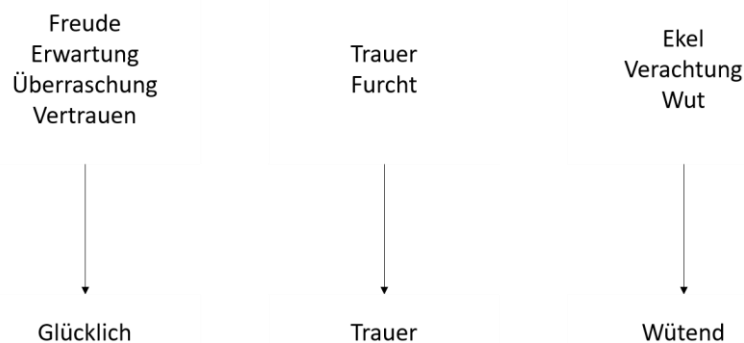


Abbildung 3: Emotionen Kategorisierung

Im Falle, dass eine eindeutige dominante Emotion nicht bestimmbar ist, beispielsweise wenn keinem Wort in dem Satz eine Emotion zugeordnet werden konnte oder zwei verschiedene Emotionen gleich oft detektiert wurde, weichen wir auf das PAD-Modell aus. Hierfür wird der Mittelwert aus den PAD-Werten aller detektierten Emotionen gebildet, gewichtet nach Anzahl ihrer Vorkommen. Da das PAD-Modell kontinuierlichen Charakters ist, erhalten wir so immer eine Emotion in PAD-Darstellung, die eine Mischung aus den detektierten Emotionen im Satz sind. Für den Spezialfall, dass keinerlei Emotionen im Satz gefunden wurde definieren wir den Satz als neutral, was einem PAD-Wert von (0, 0, 0) entspricht. Der berechnete PAD-Wert kann jetzt benutzt werden um den Satz mit der entsprechenden Emotion zu annotieren. Eine schematische Übersicht über diese Emotion-Detection-Pipeline findet sich in Abbildung 4.

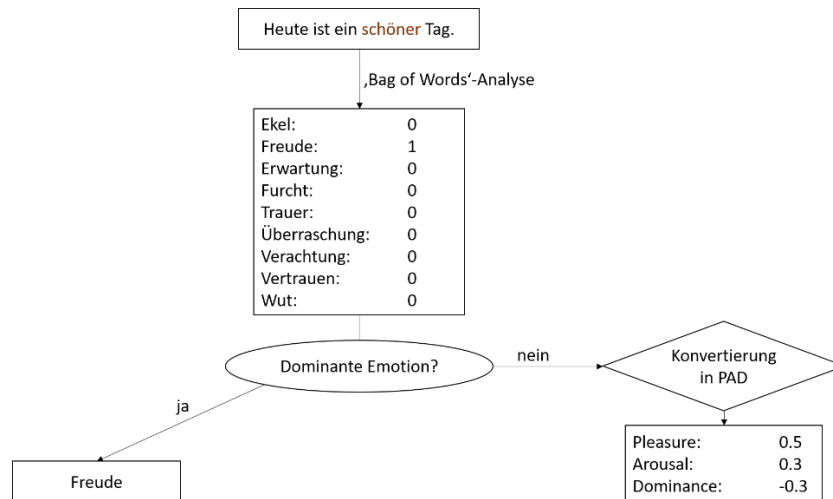


Abbildung 4: Schematische Übersicht über die Pipeline

Vor- und Nachteile

Die Stärke dieser Herangehensweise ist vor allem die sehr einfache Verarbeitung des Textes, die doch sehr solide Ergebnisse liefert. In den meisten Fällen wird eine passende dominante Emotion erkannt, welche in der von MaryTTS erstellten Audio-Datei gut hörbar umgesetzt wird. Kann diese Herangehensweise keine eindeutige Emotion im Satz erkennen, gibt es die Möglichkeit auf das PAD-Modell auszuweichen, um eine Mischung aus Emotionen zu repräsentieren. So erhält jeder Satz sicher eine passende Annotation.

Zwar ist es möglich in MaryTTS Texte auch mit PAD-Werten zu annotieren, jedoch sind Unterschiede zwischen verschiedenen PAD-Werten nach der Sprachsynthese leider kaum hörbar. Da diese Methode jeden Satz unabhängig untersucht, fließt keinerlei Kontext in die Emotionsanalyse ein. Das kann zu starken Emotions-Schwankungen zwischen Sätzen führen, was unnatürlich auf einen Zuhörer wirkt. Suboptimal ist auch, dass wir unsere Sätze nur mit jeweils 5 verschiedenen Emotionen annotieren können, obwohl es uns bei der Textanalyse möglich ist zwischen 9 verschiedenen Emotionen zu unterscheiden. Hier geht viel Potenzial für eine genaue Differenzierung zwischen Emotionen, was unsere Datenbank ermöglicht, verloren.

5.2 Ansatz mit Smoothing

Funktionsweise

Um sprunghafte Emotionswechsel zwischen aufeinanderfolgenden Sätzen zu vermeiden, müssen wir ein System in die Emotionsanalyse integrieren, welches den Kontext eines Satzes miteinbezieht. Durch das erweiterte Blickfeld bei der Analyse eines Satzes, ist es uns dann möglich plötzliche Emotionsumbrüche im Textfluss zu erkennen und entsprechend zu behandeln.

Eine sehr einfache aber effektive Methode, um die Emotionswerte vorheriger Sätze in die Analyse des aktuellen Satzes miteinzubeziehen, ist eine rekursiv angewandte Glättung der Emotionswerte. Hierbei wird der Satz zuerst unabhängig mit der ‚Bag of words‘-Analyse untersucht und der resultierende PAD-Wert mit dem PAD-Wert des vorherigen Satzes kombiniert. Die Kombination der beiden Werte lässt sich mit einem Parameter ‚*ChangeFactor*‘ $\in [0,1]$ steuern, der bestimmt, zu welchem Anteil der PAD-Wert des vorherigen Satzes gewichtet wird. Abbildung 5 verdeutlicht diesen Vorgang an einem Beispiel mit *ChangeFactor* = 0,5.

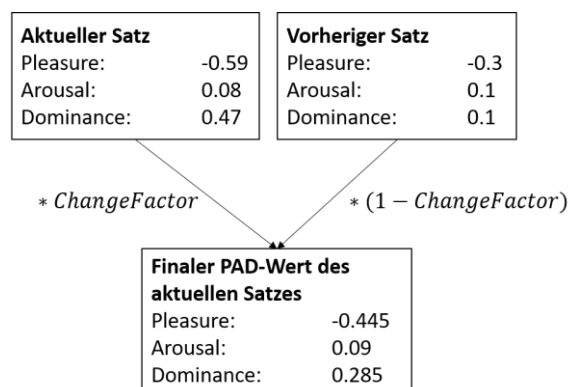


Abbildung 5: Gewichtete Kombination zweier PAD-Werte

Wir erhalten so die Möglichkeit anhand des *ChangeFactors* zu bestimmen inwiefern bei der Analyse eines Satzes die Ergebnisse vorheriger Sätze mitbeachtet werden. Für *ChangeFactor* = 1 wird jeder Satz unabhängig voneinander betrachtet, bedeutet wir haben die Möglichkeit dieses Feature einfach zu deaktivieren.

Die Kombination zweier Emotionen wird auch hier durch das PAD-Modell ermöglicht, was jedoch zur Folge hat, dass die Mischung der beiden Emotionen wieder eine Emotion als PAD-Wert liefert. Da wir vermeiden wollen den Text mit PAD-Werten zu annotieren (denn diese verursachen schlechte Ergebnisse bei der Sprachsynthese), werden bei diesem Ansatz die PAD-Werte zurück auf die konkreten Emotionen „neutral“, „glücklich“, „traurig“ und „wütend“ abgebildet. Diese konkreten Emotionen sind alle in MaryTTS darstellbar und liefern gut hörbare Emotionen.

Die Abbildung von PAD-Werte auf konkrete Emotionen erfolgt so, dass ein PAD-Wert auf eine Emotion gemappt wird, wenn der PAD-Wert der konkreten Emotion am nächsten zu

dem abzubildenden PAD-Wert ist. Als Distanzmaß wird der euklidische Abstand benutzt, welche die Differenzen der einzelnen Koordinaten des PAD-Wertes quadriert und dann aufsummiert (und anschließend die Quadratwurzel zieht, was für uns jedoch irrelevant ist, da wir nur die kleinste Distanz suchen). Dieses Distanzmaß ist durch die Charakterisierung der PAD-Werte als Punkte in einem 3-dimensionalen Raum motiviert.

Um den PAD-Wert eines Satzes zunächst zu erhalten führen wir, ähnlich wie im ersten Ansatz, zunächst eine ‚Bag of words‘-Analyse durch, die uns die Vorkommen der 9 verschiedenen Emotionen in unserer Datenbank zählt. Danach wird überprüft ob eine eindeutige dominante Emotion, also eine Emotion, die am häufigsten vorkommt, existiert. Wenn ja, dann ignorieren wir alle anderen Emotionen (setzte die Anzahl an Vorkommen auf 0) und berechnen den PAD-Wert der dominanten Emotion. Gibt es mehrere dominante Emotionen, dann berechnen wir die PAD-Werte aller dominanten Emotionen und mitteln. Beim Spezialfall, dass kein Wort im Satz mit einer Emotion in Verbindung gebracht wird, wird als PAD-Wert (0, 0, 0) angenommen.

Das Suchen nach dominanten Emotionen ist sinnvoll, da bei einer ‚Bag-of-words‘-Analyse oftmals neben einer dominanten Emotion auch andere Emotionen detektiert werden, die jedoch nicht passend den Emotionsgehalt des Satzes beschreiben. Wir nehmen deshalb einen möglichen kleinen Informationsverlust in Kauf um dieses Rauschen zu entfernen.

Ein beispielhafter Ablauf der Generierung eines PAD-Wertes ist in Abbildung 6 gegeben.

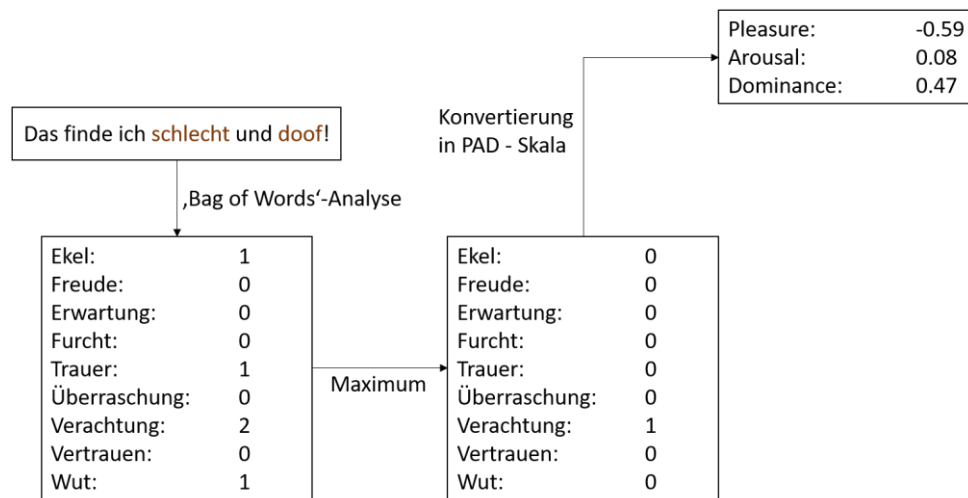


Abbildung 6: PAD-Wert eines Satzes generieren

Kombiniert man alle oben vorgestellten Bausteine erhält man den neuen Ansatz einer Emotions-Analyse-Pipeline:

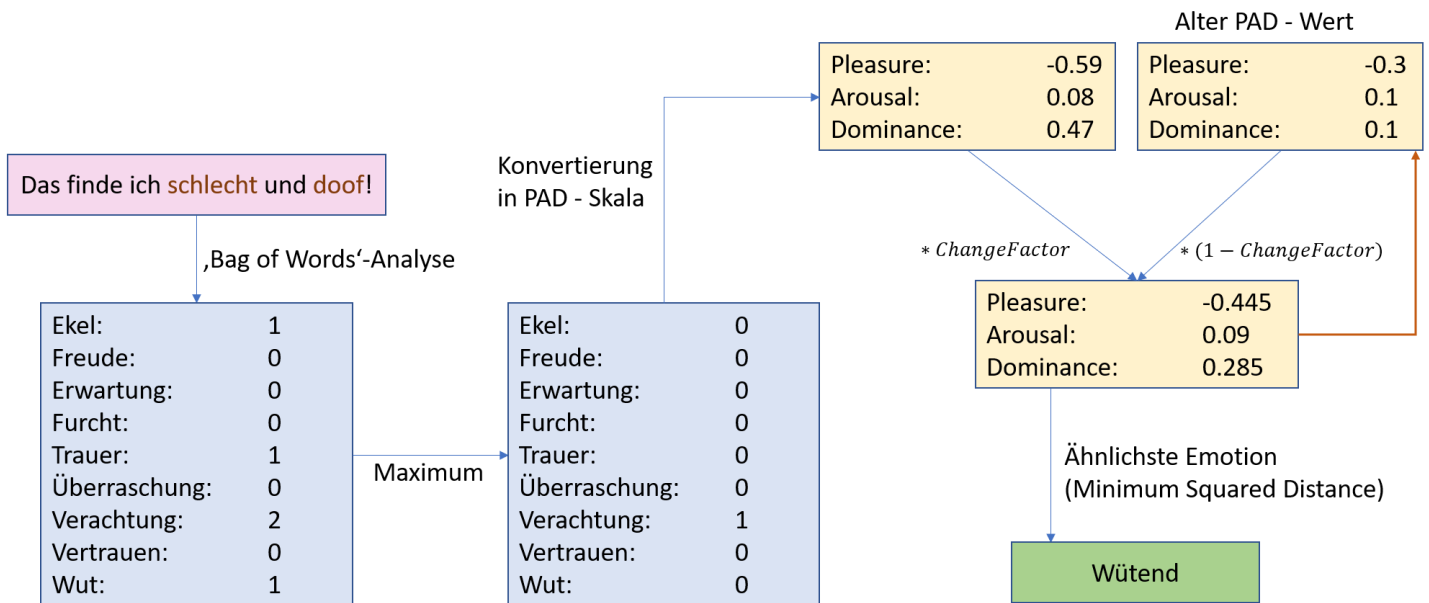


Abbildung 7: Emotions-Analyse-Pipeline mit Kontextbeachtung

Vor- und Nachteile

Im Vergleich zur ersten Methode kann man bei diesem Ansatz deutliche Vorteile erkennen: Wir haben nicht mehr das Problem, dass bei einem Satz ohne eindeutige dominante Emotion der PAD-Wert zur Annotierung verwendet werden muss, sondern immer eine eindeutige Emotion Resultat der Analyse ist. MaryTTS produziert daraus emotionale Sprache, in der Emotionen gut erkennbar sind. Trotzdem können wir die Stärken des PAD-Modells nutzen, indem wir bei der Kombination zweier Emotionen auf dem kontinuierlichen Charakter der PAD-Werte aufbauen. Auch lassen sich zu starke unnatürlich wirkende Schwankungen in den Emotionen vermeiden, indem man den *ChangeFactor* nach unten reguliert.

Ein Problem, welches jetzt neu dazukommt, ist die optimale Wahl des *ChangeFactors*. Das gesamte Ergebnis der Emotionsanalyse eines Textes hängt davon ab, weswegen wir uns im Weiteren mit dem Finden eines geeigneten Wertes beschäftigen.

6 Evaluation

Ziel der durchgeführten Evaluation ist es, die Qualität der Annotierung zu bewerten und gleichzeitig das Finden eines optimalen *ChangeFactors*, welcher besagte Qualität maximiert.

Hierfür forderten wir Testpersonen auf, eine Liste an Sätzen mit den 3 verfügbaren Emotionen ‚fröhlich‘, ‚traurig‘ und ‚wütend‘ zu annotieren, beziehungsweise keine Annotierung vorzunehmen, wenn der Satz nicht durch eine der 3 Emotionen betont werden soll.

Die selbe Liste an Sätzen wurde auch von IMMANUEL mit Emotionen annotiert, wobei wir hier mehrere Durchgänge mit unterschiedlichen *ChangeFactor* durchführten.

Das Ergebnis der Testpersonen wurde dann mit den Annotationen von IMMANUEL verglichen. Für eine objektive Beurteilung der Qualität der Annotierung nutzen wir den F-Score, ein in der Sentiment-Detection häufig verwendetes Maß für die Beurteilung eines Klassifikators.

6.1 F-Score

Der F-Score wird hauptsächlich zur Bewertung von binären Klassifikatoren benutzt. Es wird hierfür zunächst ‚*precision*‘ und ‚*recall*‘ eines Testlaufs folgendermaßen gemessen.

$$precision = \frac{true\ positives}{true\ positives + true\ negatives}$$

$$recall = \frac{true\ positives}{true\ positives + false\ positives}$$

Der F-Score kombiniert ‚*precision*‘ und ‚*recall*‘ zu einer objektiven Qualitätsbewertung:

$$F = 2 * \frac{precision * recall}{precision + recall}$$

Bei unserer Klassifikation gilt ein Testsample als ‚true positive‘, wenn die Emotion *fröhlich*, *traurig* oder *wütend* richtig erkannt worden ist. Zur Beurteilung von richtig oder falsch werden die Annotationen des Roboters mit den Resultaten der Umfrage verglichen.

Ein Testsample ist ein ‚true negative‘, wenn die Emotion *neutral* von IMMANUEL richtig erkannt worden ist. Ein ‚false positive‘ bedeutet, dass ein Sample mit einer Emotion *fröhlich*, *traurig* oder *wütend* nicht korrekt erkannt worden ist.

6.2 Resultate

Die Messung von *precision* und *recall* abhängig von dem *ChangeFactor* ergab, dass für einen *ChangeFactor* kleiner als 0,5 sowohl *precision* als auch *recall* zwischen 0 und 0,1 liegen (vergl. Abbildung 8). Das ist so zu interpretieren, dass bei einem niedrigem *ChangeFactor* und einem Text in dem die vorkommenden Emotionen schnell wechseln, der errechnete PAD-Wert um (0, 0, 0), was der Emotion ‚neutral‘ entspricht, oszilliert. Ein kleiner *ChangeFactor* verhindert, dass wenige Sätze mit der gleichen Emotion eine ausreichend große Änderung des errechneten PAD-Wertes bewirken, weshalb sich die resultierende Emotion nicht ändert. Das bewirkt eine schlechte *precision* und schlechten *recall*. Deutlich zu erkennen ist auch, dass die besten Werte für *precision* und *recall* bei einem *ChangeFactor* von 1 erzielt werden. Das spiegelt sich genauso auch im finalen F-Score wieder, welcher bei einem *ChangeFactor* von 1 maximal ist (vergl. Abbildung 9).

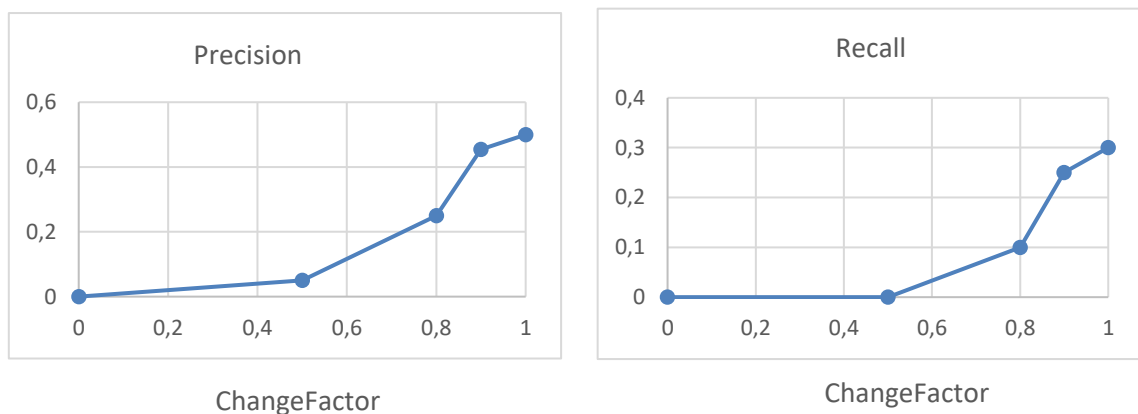


Abbildung 8: *precision* und *recall* in Abhängigkeit von *ChangeFactor*

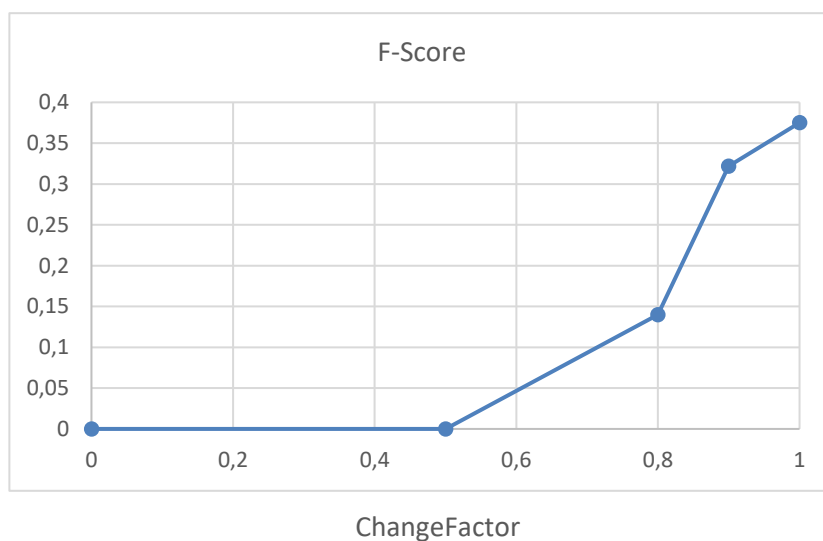


Abbildung 9: F-Score in Abhängigkeit von *ChangeFactor*

7 Diskussion

Ein *ChangeFactor* von 1 entspricht einer Emotions-Annotierung, die Informationen aus vorherigen Sätzen, bei der Untersuchung des aktuellen Satzes, nicht miteinbezieht. Jeder Satz sollte laut den Tests für bestmögliche Resultate unabhängig analysiert werden, bedeutet Smoothing sollte nur in Spezialfällen bei manueller Analyse benutzt werden.

Untersuchungen zeigen, dass Menschen eine *precision* von circa 80% erreichen, heißt 2 Menschen sind sich in 20% der Fälle uneinig welche Emotion die passende für einen Satz ist (Wilson, Wiebe & Hoffmann, 2005). Wir erreichen durch die beste Wahl des *ChangeFactors* eine *precision* von 50%, bedeutet im Vergleich zu einem menschlichen Gegenüber verhält sich IMMANUEL nur in weniger als jedem 3. Fall unnatürlich. Um noch bessere Resultate zu erzielen, könnte zukünftige Arbeit versuchen, auf der bereits kompilierten Datenbank, eine fortgeschrittenere Analyse-Methode aufzubauen. Besonders Algorithmen die auf maschinellem Lernen basieren liefern immer besserwerdende Resultate, so erreichen Prerna Chikersal, Soujanya Poria & Erik Cambria in ihrer Arbeit zur Analyse von Tweets im Jahr 2015 bereits einen F-Score von 57%.

8 Zusammenfassung

IMMANUEL ist ein Roboter, der mit Menschen diskutieren kann. Um IMMANUEL menschlicher wirken zu lassen, annotieren wir den zu sprechenden Text mit Emotionen, welche dann von der genutzten Sprachsynthesoftware umgesetzt werden. Das emotionale Befinden wird nur auf Basis des Textes automatisch generiert, ohne zusätzliche manuell beigefügte Informationen. Hierfür verwenden wir Methoden aus dem Feld der Sentiment Detection, genauer den ‚Bag of Words‘-Ansatz. Die dazu benötigte Datenbank wurde aus zwei bereits vorhandenen Datenbanken zusammengefasst und einer Stammformreduktion unterzogen, damit der Abgleich mit Wörtern aus dem Text effizient ablaufen kann. Bei unserem ersten Ansatz zur Annotation stießen wir auf das Problem, dass die genutzte Sprachsynthesoftware zwar Angaben in PAD-Werten unterstützt, diese jedoch nur schlecht hörbar umsetzt. Der zweite Ansatz umgeht dieses Problem, indem PAD-Werte wieder zurück auf konkrete Emotion abgebildet werden, welche klar und deutlich zu hören sind. Auch kommt mit dem neuen Ansatz die Möglichkeit dazu, mithilfe eines Parameters *ChangeFactor*, Informationen über Sätze hinweg zu propagieren. Eine Umfrage zur Evaluation unserer Analysemethode und zur Bestimmung eines geeigneten Wert für *ChangeFactor* hat jedoch gezeigt, dass dieses Smoothing über Sätze hinweg die Präzision unserer Analyse verschlechtert. Mit optimal gewähltem *ChangeFactor* erreicht IMMANUEL eine Präzision von 50% an richtig analysierten Emotionen, was nicht weit unter der menschlichen Leistung liegt.

Durch seine neu gewonnene emotionale Sprache wirkt IMMANUEL ein weiteres Stück menschlicher, was in denen von ihm geführten emotionalen ethisch-moralischen Diskussionen ein großer Fortschritt ist.

9 Referenzen

- Lindner, F., Bentzen, M. M., & Nebel, B. (2017). The HERA approach to morally competent robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 6991 – 6997).
- Lindner, F., & Bentzen, M. M. (2017). The hybrid ethical reasoning agent IMMANUEL, In *HRI'17 Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, Wien, Österreich (pp. 187-188).
- Martineau, J., & Finin, T. (2009). Delta TFIDF: An improved feature space for sentiment analysis. *Icwsn*, 9, 106.
- Wilson, T., Wiebe, J., & Hoffmann, P. (2005, October). Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on human language technology and empirical methods in natural language processing* (pp. 347-354). Association for Computational Linguistics.
- Pak, A., & Paroubek, P. (2010, May). Twitter as a corpus for sentiment analysis and opinion mining. In *LREc* (Vol. 10, No. 2010).
- Ogneva, M. (2010, April 19). How companies can use sentiment analysis to improve their business. *Mashable*. Abgerufen am April 10, 2018, von <https://mashable.com/2010/04/19/sentiment-analysis/#VtYgxDIgguq1>
- Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A., & Potts, C. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing* (pp. 1631-1642).
- Klinger, R., & Surayya, S. S. (2016). *Automatic Emotion Detection for Quantitative Literary Studies – A case study based on Franz Kafka's Das Schloss und Amerika*. Abgerufen von <http://www.romanklinger.de/publications/klinger-samat-reiter2016.pdf>
- Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word-emotion association lexicon. *Computational Intelligence*, 29(3), 436-465.
- Zhang, S., Wu, Z., Meng, H. M., & Cai, L. (2010). *Modeling machine emotions for realizing intelligence*. Heidelberg, Deutschland: Springer.
- Wilson, T., Wiebe, J., & Hoffmann, P. (2005, October). Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on human language technology and empirical methods in natural language processing* (pp. 347-354). Association for Computational Linguistics.
- Nasukawa, T., & Yi, J. (2003, October). Sentiment analysis: Capturing favorability using natural language processing. In *Proceedings of the 2nd international conference on Knowledge capture* (pp. 70-77). ACM.
- Chikersal, P., Poria, S., & Cambria, E. (2015). SeNTU: sentiment analysis of tweets by combining a rule-based classifier with supervised learning. In *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)* (pp. 647-651).