

Analyzing Features for the Detection of Happy Endings in German Novels

Jannidis, Fotis

fotis.jannidis@uni-wuerzburg.de
Universität Würzburg, Deutschland

Reger, Isabella

isabella.reger@uni-wuerzburg.de
Universität Würzburg, Deutschland

Zehe, Albin

zehe@informatik.uni-wuerzburg.de
Universität Würzburg, Deutschland

Becker, Martin

becker@informatik.uni-wuerzburg.de
Universität Würzburg, Deutschland

Hettinger, Lena

lena.hettinger@uni-wuerzburg.de
Universität Würzburg, Deutschland

Hotho, Andreas

hotho@informatik.uni-wuerzburg.de
Universität Würzburg, Deutschland

Einleitung

Note: An English version of this paper is available from <https://arxiv.org/abs/1611.09028>.

Der Plot ist ein grundlegendes Strukturelement literarischer Texte. Dementsprechend wären Methoden zur computergestützten Repräsentation von Plot oder bestimmten Plot-Elementen ein großer Gewinn für die quantitative Literaturanalyse. Dieses Paper betrachtet ein solches Plot-Element: das Ende; genauer gesagt untersuchen wir die Frage, ob ein Werk ein Happy End hat oder nicht. Dazu setzen wir Sentimentanalyse ein, wobei wir den Fokus auf die qualitative Betrachtung bestimmter Features und deren Performanz legen, um tiefere Einsicht in die Funktionsweise der automatischen Klassifikation zu erhalten. Außerdem zeigen wir, wie die beschriebene Vorgehensweise auf nachfolgende Forschungsfragen angewendet werden und dabei zu interessanten Ergebnissen hinsichtlich der Erscheinungszeit der Romane führen kann.

Verwandte Arbeiten

In einer der ersten Arbeiten beschäftigt sich Mark Finlayson mit folkloristischen Erzählungen und entwickelt einen Algorithmus, der Ereignisse erkennt und daraus übergeordnete Konzepte wie Niedertracht oder Belohnung abstrahiert (Finlayson 2012). Reiter et al. identifizieren Ereignisse sowie deren Teilnehmer und Reihenfolge und nutzen maschinelle Lernverfahren, um strukturelle Ähnlichkeiten über Erzählungen hinweg aufzudecken (Reiter 2013, Reiter et al. 2014).

In letzter Zeit richtet sich einige Aufmerksamkeit auf die Sentimentanalyse, insbesondere seit Matthew Jockers emotionale Erregung als Indikator für Plotstrukturen vorgeschlagen hat (Jockers 2014). Er unterteilt Romane in Segmente und bildet daraus emotionale Plot-Kurven (Jockers 2015). Obwohl die Idee, Sentimentanalyse in diesem Zusammenhang einzusetzen, gut aufgenommen wurde, wurde Jockers für seine Verwendung der Fourier-Transformation zur Glättung der resultierenden Plot-Kurven kritisiert (Swafford 2015, Schmidt 2015).

Micha Elsner (Elsner 2015) verwendet, neben anderen Features, ebenfalls Sentimentkurven, um Repräsentationen des Plots romantischer Werke zu erstellen. Er verknüpft diese Kurven mit bestimmten Figuren und untersucht auch das gemeinsame Auftreten von Figuren. Die Auswertung seines Ansatzes zeigt, dass er echte Romane mit beachtlichem Erfolg von künstlich umgestellten Versionen unterscheiden kann, was darauf hindeutet, dass seine Methoden tatsächlich bestimmte Aspekte der Plotstruktur abbilden.

In vorhergehenden Arbeiten haben wir Sentiment-Features verwendet, um Happy Ends, als ein wichtiges Plot-Element, in deutschsprachigen Romanen zu erkennen, wobei wir einen F1-score von 73% erreichen konnten (Zehe et al. 2016).

Korpus und Ressourcen

Unser Datensatz besteht aus 212 deutschsprachigen Romanen, die hauptsächlich aus dem 19. Jahrhundert stammen. Zu jedem Roman wurde manuell annotiert, ob er ein Happy End hat (50%) oder nicht (50%). Die dafür relevanten Informationen stammen aus den Zusammenfassungen des Kindler Literatur Lexikon Online und aus Wikipedia. Sofern keine Zusammenfassung eines Romans verfügbar war, wurde das Ende von den Annotatoren gelesen.

Unsere Sentimentanalyse erfordert eine Ressource, die auflistet, welche Gefühle Leser typischerweise mit bestimmten Worten oder Phrasen eines Textes assoziieren. Dieses Paper verwendet das NRC Sentiment Lexikon (Mohammad und Turney 2013), zu dem eine automatisch übersetzte deutsche Version verfügbar ist. Eine besondere Eigenschaft dieses Lexikons ist, dass zu jedem Wort neben je einem binären Wert (0 oder 1) für positive und negative

Konnotation (2 Features) auch seine Zugehörigkeit zu 8 Basisemotionen (Wut, Angst, Ekel, Überraschung, Freude, Vorfreude, Vertrauen und Trauer) festgehalten ist (vgl. Tabelle 1). Zusätzlich ermitteln wir die Polarität eines Wortes, indem der negative vom positiven Wert abgezogen wird (ein Wort mit einem positiven Wert von 0 und einem negativen Wert von 1 erhält also die Polarität -1). Die Polarität dient als ein zusammengefasster Emotionswert. Insgesamt betrachten wir also 11 Features.

Tabelle 1: Beispieleinträge aus dem NRC Sentiment Lexikon

Wort/ Dimension	verabscheuen	bewundernswert	Zufall
Positiv	0	1	0
Negativ	1	0	0
Polarität	-1	1	0
Wut	1	0	0
Vorfreude	0	0	0
Ekel	1	0	0
Angst	1	0	0
Freude	0	1	0
Trauer	0	0	0
Überraschung	0	0	1
Vertrauen	0	1	0

Experimente

Ziel dieses Papers ist es, Features, die zur Erkennung von Happy Ends in Romanen genutzt wurden, genauer zu untersuchen, um Einsichten in die Relevanz bestimmter Features zu erhalten. Dazu übernehmen wir die Features und Methoden, wie sie in Zehe et al. (2016) beschrieben sind. Die Parameter der linearen SVM sowie die Einteilung in 75 Segmente sind ebenfalls aus diesem Paper übernommen.

Features. Da keine verlässlichen Kapitelannotationen verfügbar waren, wurde jeder Roman in 75 gleichgroße Blöcke unterteilt, die wir als *Segmente* bezeichnen. Für jedes lemmatisierte Wort werden die oben beschriebenen 11 Sentiment-Werte ermittelt. Anschließend wird für jedes Segment der entsprechende Durchschnitt berechnet, sodass 11 Werte pro Segment vorliegen. Diese werden als ein Feature-Set betrachtet.

Qualitative Feature-Analyse. Da unser Korpus zu gleichen Teilen aus Romanen mit und ohne Happy End besteht, erreichen sowohl die Random Baseline, als auch die Mehrheits-Baseline eine Klassifikationsgenauigkeit von 50%.

Aufgrund unserer Annahme, dass die relevante Information zur Klassifikation von Happy Ends am Ende eines Romans zu finden ist, wurden zunächst die Sentiment-Werte des letzten Segments als einziges Feature-

Set ($f_{d, n}$) verwendet, was zu einer Genauigkeit von 67% führte.

Um unserer Intuition gerecht zu werden, dass nicht nur das letzte Segment an sich, sondern auch sein Verhältnis zum Rest des Romans für die Klassifikation von Bedeutung ist, wurden sogenannte Sektionen (*sections*) eingeführt: das letzte Segment eines Romans bildet die *final section*, während die übrigen Segmente zur *main section* gehören. Über die Sektionen wurden wiederum Durchschnittswerte gebildet, indem der jeweilige Wert aller 11 Features über alle Segmente in der betreffenden Sektion gemittelt wurde. Um das Verhältnis zwischen diesen Sektionen abzubilden, wurden die Differenzen zwischen den Sentiment-Werten der final section und den durchschnittlichen Sentiment-Werten aller Segmente in der main section als zusätzliche Features betrachtet. Dies hatte jedoch keine Auswirkungen auf die Ergebnisse.

Diese Beobachtung führte uns zu der Annahme, dass unser Begriff des "Endes" nicht differenziert genug ist, da die Anzahl an Segmenten für jeden Roman und damit auch die Grenzen des finalen Segments relativ willkürlich gewählt wurden. Daher wurde die Aufteilung in final section und main section im Folgenden variiert, sodass die final section mehr als nur das letzte Segment enthalten kann.

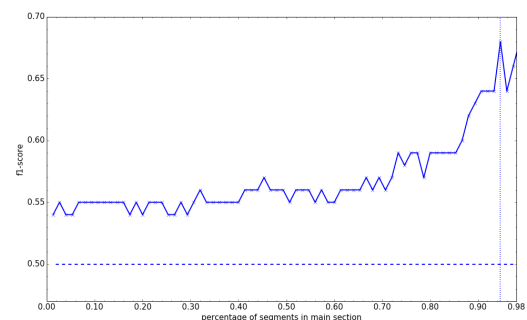


Abbildung 1: Klassifikationsgenauigkeit für verschiedene Unterteilungen in main und final section. Die gestrichelte Linie gibt die Baseline an, die gepunktete Linie markiert die Aufteilung, bei der der maximale F1-score erreicht wird.

Abbildung 1 zeigt, dass die Klassifikationsgenauigkeit steigt, wenn mindestens 75% der Segmente in der main section sind und ein Maximum bei ca. 95% erreicht (bei 75 Segmenten insgesamt bedeutet das 4 Segmente in der final section und 71 Segmente in der main section). Mit dieser Aufteilung verbessert sich der F1-Wert auf 68%, wenn nur das Feature-Set der final section ($f_{d, final}$) verwendet wird, und weiter auf 69%, wenn die Differenzen zu den durchschnittlichen Sentiment-Werten der main section ($f_{d, main - final}$) miteinbezogen werden.

Da sich die Ergebnisse durch die Einbeziehung des Verhältnisses zwischen der final section und der main section verbessert haben, war unser nächster Schritt, den Verlauf der Sentimentkurve gegen Ende eines Romans genauer zu modellieren. Beispielsweise könnte sich kurz vor dem Ende eine Katastrophe ereignen, die anschließend im Sinne eines Happy Ends aufgelöst wird. Um diese Intuition abzubilden, führten wir eine weitere Sektion ein, die sogenannte late-main section, die die letzten Segmente der main section umfasst. Die Differenzen zwischen den Feature-Sets für die late-main section und die final section wurden als zusätzliche Merkmale verwendet ($f_d, \text{late} - \text{final}$). Mit diesen 3 Feature-Sets erzielten wir einen F1-score von 70%. Durch die zusätzliche Verwendung des Feature-Sets für das letzte Segment stieg der F1-score auf 73%.

Tabelle 2: F1-score für die verschiedenen Feature-Sets

Features	Ergebnisse
1) Feature-Set finales Segment	67%
2) Feature-Set finales Segment und Differenz zur main section	67%
3) Feature-Set final section mit final section der Länge 4	68%
4) Feature-Set 3 und Differenz zur main section	69%
5) Feature-Set 4 und Differenz zwischen late-main section und final section	70%
6) Feature-Set 5 und Feature-Set finales Segment	73%

Die beschriebenen Ergebnisse sind in Tabelle 2 zusammengefasst. Hier wird deutlich, dass die Aufnahme der einzelnen Feature-Sets jeweils zu einer kleinen Verbesserung geführt hat, bis hin zu einem F1-score von 73%. Obwohl die Aufteilung mit 4 Segmenten in der final section die besten Ergebnisse erzielte, konnten wir auch beobachten, dass einige Romane mit mehreren verschiedenen Unterteilungen korrekt klassifiziert werden konnten. Andere Romane hingegen konnten in keinem Setting korrekt vorhergesagt werden. Als Beispiel sei hier Jules Vernes Roman *Zwanzigtausend Meilen unter dem Meer* genannt, der ein eindeutiges Happy End mit klaren Grenzen hat, das jedoch extrem kurz ist und nur aus den ca. 250 letzten Wörtern besteht. Diese Beobachtungen zeigen, dass der Begriff des "Endes" eines Roman sehr variabel ist und von Text zu Text sehr unterschiedlich manifestiert sein kann.

Korrelation mit Erscheinungszeit. Das wirft wiederum die Frage auf, ob die Sensibilität unserer Methode hinsichtlich solcher Variabilität genutzt werden kann, um gewisse Eigenschaften der Romane in unserem Korpus besser zu verstehen. Als Beispiel haben wir untersucht, ob und inwiefern der Erfolg verschiedener Unterteilungen von Romanen mit deren Erscheinungsdatum korrelieren. Um die Ergebnisse so gut wie möglich interpretierbar

zu halten, beschränken wir uns auf ein Feature-Set: die Sentiment-Werte der finalen Sektion. Zunächst haben wir unser Korpus in 4 Gruppen unterteilt: Romane, die vor 1830 erschienen sind (65 Texte), zwischen 1831 und 1848 (31 Texte), zwischen 1849 und 1870 (29 Texte) und nach 1871 (87 Texte). Diese Einteilung führte zu ähnlich großen Untergruppen, von denen keine eine besondere Tendenz hinsichtlich Romanen mit oder ohne Happy End aufweist.

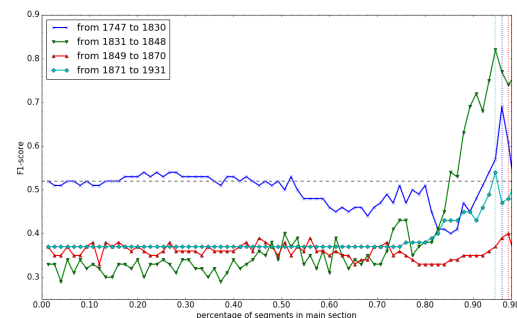


Abbildung 2: F1-score für verschiedene Unterteilungen in main und final section. Die farbigen Kurven stehen für Romane aus verschiedenen Zeitperioden. Die gestrichelte Linie zeigt die Zufallsbaseline für die Zeitperiode ab 1871. Die Baselines für die anderen Zeitperioden liegen etwas darunter und werden daher nicht dargestellt. Die gepunkteten Linien zeigen jeweils den maximalen F1-Wert für die entsprechende Zeitperiode.

Abbildung 2 zeigt, dass die Klassifikation erneut dann am besten funktioniert, wenn ca. 95-98% der Segmente in der Hauptsektion sind, unabhängig von der Zeitperiode. Die beste Aufteilung in Sektionen korreliert also nicht mit dem Erscheinungsjahr eines Romans. Es fällt jedoch auf, dass die Romane nach 1848 deutlich niedrigere Werte liefern als die vor diesem Jahr veröffentlichten Texte, meistens sogar unterhalb der Baseline. Das deutet auf eine Korrelation zwischen dem Erscheinungsdatum und der Klassifikationsgenauigkeit hin: Vor dem Realismus erschienene Romane sind hinsichtlich des Happy Ends leichter zu klassifizieren als realistische Romane. Eine mögliche Erklärung für diese Beobachtung könnte die stärker schematische Struktur der vor-realistischen Romane sein.

Wir sind uns bewusst, dass die Anzahl der Romane für die einzelnen Zeitperioden relativ klein ist, sodass diese Beobachtungen zunächst als exploratorische Einblicke gesehen werden müssen. Nichtsdestotrotz zeigen diese vorläufigen Ergebnisse, dass die automatische Erkennung von Happy Ends, sogar mit nur einem recht einfachen Feature-Set, Zusammenhänge zu anderen Eigenschaften von Romanen aufdecken kann, die für die Literaturwissenschaft von großem Interesse sind.

Fazit und zukünftige Arbeiten

Die automatische Erkennung von Happy Ends als wesentlichem Plot-Element von Romanen ist ein nützlicher Schritt in Richtung einer umfassenden computergestützten Repräsentation des Plots literarischer Texte. Unsere Experimente zeigen, dass verschiedene Features auf Basis von Sentimentanalyse eine Erkennung von Happy Ends in Romanen mit unterschiedlicher, aber insgesamt solider Genauigkeit ermöglichen. Obwohl unser Ansatz relativ einfach gehalten ist, kann er zu substantiellen Erkenntnissen für die Literaturwissenschaft führen.

In zukünftigen Arbeiten soll die Genauigkeit unserer Methode verbessert werden, indem die hohe Variabilität des Endes in Romanen differenzierter betrachtet wird. Außerdem könnte der Ansatz eingesetzt werden, um bestimmte Eigenschaften weiterer Romankorpora tiefergehend zu untersuchen.

Fußnoten

1. Quelle: <https://textgrid.de/digitale-bibliothek>
2. www.kll-online.de
3. <https://de.wikipedia.org>
4. <http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>

Bibliography

Elsner, Micha (2015): „Abstract Representations of Plot Structure“, in: *Linguistic Issues in Language Technology* 12 (5).

Finlayson, Mark A. (2012): *Learning Narrative Structure from Annotated Folktales*. PhD thesis, Massachusetts Institute of Technology.

Jockers, Matthew L. (2014): *A novel method for detecting plot*. <http://www.matthewjockers.net/2014/06/05/a-novel-method-for-detecting-plot/> [letzter Zugriff 25. August 2016].

Jockers, Matthew L. (2015): *The rest of the story*. <http://www.matthewjockers.net/2015/02/25/the-rest-of-the-story/> [letzter Zugriff 25. August 2016].

Mohammad, Saif / Turney, Peter (2013): „Crowdsourcing a Word-Emotion Association Lexicon“, in: *Computational Intelligence* 29 (3): 436–465.

Reiter, Nils (2013): *Discovering Structural Similarities in Narrative Texts using Event Alignment Algorithms*. PhD thesis, Heidelberg University.

Reiter, Nils / Frank, Anette / Hellwig, Oliver (2014): „An NLP-based Cross-Document Approach to Narrative Structure Discovery“, in: *Literary and Linguistic Computing* 29 (4): 583–605 10.1093/lc/fqu055.

Schmidt, Benjamin M. (2015): *Commodius vici of recirculation: the real problem with Syuzhet*. [http://benschmidt.org/2015/04/03/commodius-](http://benschmidt.org/2015/04/03/commodius-vici-of-recirculation-the-real-problem-with-syuzhet/)

[vici-of-recirculation-the-real-problem-with-syuzhet/](http://benschmidt.org/2015/04/03/commodius-vici-of-recirculation-the-real-problem-with-syuzhet/) [letzter Zugriff 25. August 2016].

Swafford, Annie (2015): „Problems with the Syuzhet Package“. <https://annieswafford.wordpress.com/2015/03/02/syuzhet/> [letzter Zugriff 25. August 2016].

Zehe, Albin / Becker, Martin / Hettinger, Lena / Hotho, Andreas / Reger, Isabella / Jannidis, Fotis (2016): „Prediction of Happy Endings in German Novels“, in: *Proceedings of the Workshop on Interactions between Data Mining and Natural Language Processing 2016*.