



Fakultät Informatik

Automatische Generierung von Themenspezifischen Podcasts aus einer großen Audiodatenbank

Bachelorarbeit im Studiengang Medieninformatik

vorgelegt von

Vincent Neumann

Matrikelnummer 358 5287

Erstgutachter: Prof. Dr. Florian Gallwitz

Zweitgutachter: Prof. Dr. Anna Kruspe

© 2023

Inhaltsverzeichnis

Textzusammenfassung	iii
Priorisierung der Literatur	iv
Anhang	v
Literaturverzeichnis	vi

Textzusammenfassung

In dem Artikel „NewsPod: Automatic and Interactive News Podcasts“ [Laba 22] wird ein neuer Ansatz für die automatische Erstellung von Podcast Episoden vorgestellt. Dabei ermitteln mehrere Machine Learning Systeme den Inhalt von bestimmten Nachrichtenseiten und extrahieren daraus Fragestellungen, die den Inhalt des Textes widerspiegeln sollen. Die Autoren benutzen ein GPT-2 Sprachmodell, das auf Fragengenerierung trainiert wurde, um aus jedem Absatz 7 Fragen zu ermitteln. Für jede Frage wird dann eine Separate Antwort generiert. Der Wichtigste Aspekt an diesem Artikel ist, dass die Benutzer der Software Interaktiv agieren können und selbstgewählte Fragen mithilfe eines Mikrofons oder einer Tastatur stellen können, die dann von dem System beantwortet werden. Die Autoren dieses Papers führen außerdem zwei Studien zur Nutzung dieses Systems durch. Der Gegenstand der ersten Studie ist das die Zufriedenheit einer Testgruppe mit der Erzählweise des Textes und der automatisch generierten Stimme des Sprechers. Die zweite Studie untersucht die Interaktion der Zuhörer während der Benutzung des Systems. [Laba 22]

In der ersten Studie konnten die Autoren feststellen, dass einer ihrer Ansätze, QA Best, so erfolgreich war, dass „80% of QA Best listeners said they would use the system to listen to the news in the future“ [Laba 22]. Die zweite Studie kommt zu dem Schluss, dass zwar die Bereitschaft eigene Fragen zu stellen mit 85% der Zuhörer sehr hoch war, die Zufriedenheit der Tester mit der Qualität der Antworten aber sehr gering ausfiel, da „76% of the answers were rated Irrelevant/Confusing“ [Laba 22].

Priorisierung der Literatur

Die wichtigste Literatur stellt der Artikel von [Laba 22] dar, da er erst vor einem Jahr erschienen ist und sich auch mit der automatischen Generierung von Podcasts befasst. [Laba 22] beschäftigt sich mit der Extrahierung von Semantischen Informationen aus Nachrichtenberichten und der Aufbereitung dieser in Form von sprachsynthetisierten Podcast Episoden. Dadurch wird dieser Artikel sehr relevant für meine eigene Bachelorarbeit werden.

Als zweites würde ich die Arbeit von [Zhan 20] priorisieren, da sie als Vorarbeit von [Laba 22] dient und das wichtige Summarization-Modell PEGASUS vorstellt, das ich auch in meiner Arbeit verwenden kann.

Als weitere Quellen werde ich [Karp 20], [Redd 19] und [Choi 18] verwenden, die sich alle mit der automatischen Fragen- und Antwortgenerierung aus Texten beschäftigen. [Karp 20] ist dabei der neueste Ansatz zur Ermittlung von Kontext, der sich außerdem von dem TF-IDF Algorithmus unterscheidet und sehr gute Ergebnisse verspricht.

Für das Design der Podcast Episoden werde ich die Artikel von Podcast Übersicht von [Jones 21] verwenden, in dem mithilfe von Deep Learning Ansätzen mehr als 100.000 Podcast Episoden analysiert wurden und daraus Aspekte für eine erfolgreiche Podcast Episode kristallisiert wurden. Diese Aspekte werden mir bei der Gestaltung und dem Aufbau einer Episode helfen.

In dem Paper [Kang 12] untersuchten die Autoren den Effekt von mehreren Stimmen auf das Podcasterlebnis. Außerdem untersuchten Sie in Ihrer Studie die Effekte des Erzählstils und finden heraus, dass informelle Sprache und ein ungezwungener Kommunikationsstil das Erlebnis von Zuhörern verbessert.

Im weiteren werde ich noch die Arbeiten von [Maro 20], [Clar 20] und [Du 17] verwenden, die sich mit Deep Learning Modellen zur Erschließung von Texten beschäftigen.

Anhang

Nein ich habe keine KI-Tools verwendet um dieses Aufgabe zu lösen. KI Tools haben die Möglichkeit schnell und einfach qualitativ hochwertige Texte zu erstellen. Leider können solche Tools auch falschinformationen halluzinieren, die aber auf den ersten Blick richtig aussehen.

Literaturverzeichnis

- [Choi 18] E. Choi, H. He, M. Iyyer, M. Yatskar, W.-t. Yih, Y. Choi, P. Liang, and L. Zettlemoyer. “QuAC : Question Answering in Context”. Aug. 2018.
- [Clar 20] K. Clark, M.-T. Luong, Q. V. Le, and C. D. Manning. “ELECTRA: Pre-training Text Encoders as Discriminators Rather Than Generators”. March 2020.
- [Du 17] X. Du, J. Shao, and C. Cardie. “Learning to Ask: Neural Question Generation for Reading Comprehension”. Apr. 2017.
- [Jones 21] R. Jones, B. Carterette, A. Clifton, M. Eskevich, G. J. F. Jones, J. Karlgren, A. Pappu, S. Reddy, and Y. Yu. “TREC 2020 Podcasts Track Overview”. March 2021.
- [Kang 12] M. Kang and U. Gretzel. “Effects of Podcast Tours on Tourist Experiences in a National Park”. *Tourism Management*, Vol. 33, No. 2, pp. 440–455, Apr. 2012.
- [Karp 20] V. Karpukhin, B. Oğuz, S. Min, P. Lewis, L. Wu, S. Edunov, D. Chen, and W.-t. Yih. “Dense Passage Retrieval for Open-Domain Question Answering”. Sep. 2020.
- [Laba 22] P. Laban, E. Ye, S. Korlakunta, J. Canny, and M. Hearst. “NewsPod: Automatic and Interactive News Podcasts”. In: *27th International Conference on Intelligent User Interfaces*, pp. 691–706, ACM, Helsinki Finland, March 2022.
- [Maro 20] D. Maroni, J. Köhler, J. Fisseler, and S. Becker. “KÜNSTLICHE INTELLIGENZ IM PRODUKTIVEN EINSATZ FÜR DIE AUTOMATISIERTE ERSCHLIESSUNG IM MULTIMEDIALEN PRODUKTIONSPROZESS”. 2020.
- [Redd 19] S. Reddy, D. Chen, and C. D. Manning. “CoQA: A Conversational Question Answering Challenge”. *Transactions of the Association for Computational Linguistics*, Vol. 7, pp. 249–266, Nov. 2019.
- [Zhan 20] J. Zhang, Y. Zhao, M. Saleh, and P. Liu. “PEGASUS”. In: *Proceedings of the 37th International Conference on Machine Learning*, pp. 11328–11339, PMLR, Nov. 2020.