


Chapter 13

Enhancing Spoken Text With Punctuation Prediction Using N-Gram Language Model in Intelligent Technical Text Processing Software

Shweta Rani

 <https://orcid.org/0000-0002-9285-7073>

Jaypee Institute of Information Technology, India

Rhea Jain

Jaypee Institute of Information Technology, India

ABSTRACT

Communication is a very important practice between two individuals, and for effective communication, the spoken text must be understood by others. Punctuation prediction is utmost essential in spoken text for bridging the language gaps. Various techniques have been proposed in the literature and are also explored. In this work, the authors developed software by studying n-gram model with probability to restore the punctuation in spoken text of technical lectures. In this chapter, the authors compared uni-gram, bigram, trigram, and quadgram method on varying size of datasets. Findings suggest that trigram model outperform the other for all three datasets and it was also noticed that increasing the gram size more do not have much impact on the performance of the software.

1. INTRODUCTION

Generally, automatic speech recognition (ASR) system produces the plain text without punctuation marks. This type of text is not easy to read and may lead to different meaning if not understood carefully. However, for effective communication, punctuation marks play vital role. It plays a vital role in

DOI: 10.4018/979-8-3693-3502-4.ch013

bridging language gaps, facilitating knowledge exchange, and empowering individuals to express their ideas clearly. In today's diverse and interconnected world, ensuring that everyone has access to accurate communication is paramount. By adding the punctuation marks or the missing information in speech recognition system's translated text, it increases the understandability of the text as well as the text can be used for further use (Jones et al., 2005; Huang & Zweig, 2002; Makhoul et al., 2005).

In prior studies, researchers (Peitz et al., 2011; Cattoni et al., 2007) have shown that for machine translation, punctuation marks are crucial. (Abhishek, 2021) also conducted a useful survey on the role of futuristic algorithms. Usually, written English is organized and punctuation marks can be added with less efforts. However, the scenario is quite complicated in case of spoken English. It may contain grammatical mistakes, long pauses even not required. One may change the sentence while speaking, so these might be the challenges for punctuation restoration for spoken text. For spoken marks, various punctuation marks might be added like full stop/ period (.), semicolon (;), question mark (?), exclamation mark (!), comma (,).

This chapter delves into the pivotal realm of punctuation prediction, harnessing the power of n-grams to enhance its accuracy and, in turn, elevate the performance of language models for our objective. This research chapter seeks to investigate and explore the effectiveness of various N-gram models in accurately predicting punctuation across diverse linguistic contexts and domains..

The primary contribution of this chapter are outlined as follows:

- We developed a text processing software that is easy to use by the user for checking the punctuation marks in technical text.
- We trained the model on spoken text of technical content in engineering field.
- Training of the model is repeated multiple times to maintain and improve the efficacy of the model.
- We investigated the impact of different gram, $n=1$ to 4, how it impacts the correctness of the software.
- We compared the performance on different dataset with different size and the punctuation marks available.

2. LITERATURE REVIEW

Punctuation prediction holds significant importance in natural language processing as it enhances text comprehension and understanding. Over the years, several studies have been conducted to explore different approaches and techniques for accurate punctuation prediction. This section discusses the literature work done in the field of punctuation prediction. Table 1 lists the related work in this area.

Initially, in 1996, n gram based language model was suggested for text segmentation (Stolcke & Shriberg, 1996). They only worked upon sentence boundary. Agbago et al. (2005) proposed scoring function involving casing probability for enhancing n-gram model based on word classes for unknown words.

(Beeferman et al, 1998) implemented a trigram model for punctuation prediction. As per this study, only 14.7% trigrams were found containing commas. Later, (Kaufmann and Kalita, 2010) also utilized a trigram model for capitalization in Twitter messages. They identified the problem that tweets may not contain the correct capitalized words and the same can be corrected by n gram model.

15 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/enhancing-spoken-text-with-punctuation-prediction-using-n-gram-language-model-in-intelligent-technical-text-processing-software/346332

Related Content

The Human Experience of Algorithmic HR: A Design-Sensitive, Semi-Structured Review

Ian Mizuki and Zoya Khan (2026). *Employee-AI Experience in Human Resource Management* (pp. 35-78).

www.irma-international.org/chapter/the-human-experience-of-algorithmic-hr/413049

A Secure Remote User Authentication Protocol for Healthcare Monitoring Using Wireless Medical Sensor Networks

Preeti Chandrakar (2019). *International Journal of Ambient Computing and Intelligence* (pp. 96-116).

www.irma-international.org/article/a-secure-remote-user-authentication-protocol-for-healthcare-monitoring-using-wireless-medical-sensor-networks/216472

Sounds Relaxing—Looks Cool: Audio and Visual Selections for Computer Systems that Support Wellness

Stuart Cunningham and Rich Picking (2012). *International Journal of Ambient Computing and Intelligence* (pp. 40-53).

www.irma-international.org/article/sounds-relaxing-looks-cool/64190

Zalando's Use of AI to Create a Personalized Shopping Experience

Muntazir Mahdi (2026). *Customer-Centric AI: Conversational Technologies, Personalization, and Ethical Innovation* (pp. 145-174).

www.irma-international.org/chapter/zalandos-use-of-ai-to-create-a-personalized-shopping-experience/388521

GanglioNav WithYou: Design and Implementation of an Artificial Intelligence-Enabled Cognitive Assessment Application for Alzheimer's Patients

Vinu Sherimon, Sherimon Puliprathu Cherian, Rahul V. Nair, Khalid Shaikh and Natasha Renchi Mathew (2023). *Handbook of Research on Advancements in AI and IoT Convergence Technologies* (pp. 314-329).

www.irma-international.org/chapter/ganglionav-withyou/330073