MULTIDISCIPLINARY SCIENTIFIC JOURNAL
NÚCLEO DO CONHECIMENTO
REVISTA CIENTÍFICA MULTIDISCIPLINAR NÚCLEO DO CONHECIMENTO ISSN: 2448-0959
https://www.nucleodoconhecimento.com.br

USE OF ARTIFICIAL INTELLIGENCE IN BIBLICAL CITATION RECOMMENDATIONS IN THE NEW TESTAMENT

ORIGINAL ARTICLE

LIMA, Bruno Cesar Dos Santos[1], OMAR, Nizam[2], AVANSI, Israel[3], CASTRO, Leandro Nunes de[4], SILVEIRA, Ismar Frango[5]

RIBEIRO, Viviane Pinto Alves. **Use of artificial intelligence in biblical citation recommendations in the New Testament**. Revista Científica Multidisciplinar Núcleo do Conhecimento. Year. 08, Ed. 07, Vol. 02, pp. 123-143. Julho 2023. ISSN: 2448-0959, Access link: https://www.nucleodoconhecimento.com.br/computer-engineering/biblical-citation

ABSTRACT

Religion occupies a prominent place in people's daily lives and is made explicit to the public or the faithful through preaching or exposition of their sacred texts. The Holy Bible is the religious literature of Christianity, and its text has a unique nature of interpretation and knowledge extraction, that is, through the reading done by specialists (theologians). However, an automated knowledge extraction or that involves some automatic mechanism intelligence to support the interpretation (hermeneutics) of the Biblical text is not observed in the literature. Probably this gap in the literature is caused by the complexity of the biblical textual corpus and the multiplicity of genres it has, being an interpretative challenge even for human specialists. Therefore, this article primarily seeks to build an automated way through artificial intelligence (AI) to provide contextual biblical quotations from the four gospels of the New Testament for the construction of sermons or development of homiletics, which is the art of producing religious sermons for teaching and interpretation of the Biblical message. The methodology used in this article seeks to employ artificial intelligence techniques to implement the proposed solution, that is, a hybrid recommendation system to quote texts from Biblical passages. The AI techniques involved are text mining, natural language processing and supervised learning. Secondarily, this work aims to verify whether the combination of natural language processing techniques and machine learning can provide subsidies for the recovery or extraction of knowledge from complex textual corpus analogous to the biblical corpus. The results show that the proposed hybrid recommendation system is capable of extracting semantic and contextual meaning from the Biblical text, fundamental in the construction of homiletics. The performance evaluation metrics indicate the robustness of the results and consequently validate the findings of this research. Therefore, the

combination of these techniques can be extrapolated by the scientific community to aid in the interpretive recovery of complex textual corpus.

Keywords: Information Retrieval, Text mining, Recommendation System, Holy Bible.

1. INTRODUCTION

Religion has been a protagonist in the history of human civilization for millennia and in recent centuries its impact on society is immense, from: benefits to people's mental health to the political development of countries and Christianity within this religious phenomenon stands out among the other movements as the most influential in terms of scope and numbers, corresponding to about 31% of the world's population declaring themselves Christian[6]. (ATEN, 2019; SEN, 2022) (LUTZ, 1984)

The religious literature adopted in Christianity is the Holy Bible, and this literature is the most published in the world reaching the incredible mark of 5 billion of copies published[7] It is worth noting that biblical literature is complex and difficult to exegesis, due to its nature of multiplicity of genres and variety of linguistic phenomena (CHAPMAN, 2016) (ROGERSON, 2006). Therefore, the textual interpretation of the holy bible becomes a challenge for humans, perhaps this justifies the absence of automated techniques for biblical exegesis in the academic literature and adding to this problem the fact that for machines human language is not trivial. (MITKOV, 2003).

The way used by religions to disseminate their message is through oratory and preaching. Especially this phenomenon is observed by Christianity and this concept is found within the biblical literature itself "So then faith comes by hearing, and hearing by the word of God." - Romans 10:17, therefore, good communication is essential. Speakers generally use homiletics to produce their oral or written messages, however, it is a meticulous job that requires some experience and knowledge from the speaker to conduct it. It is possible to infer that the process of building sermons for the communication of a religious message can demand time, skill and that automated forms could help in this homiletic task.

RC: 146578
Available in: https://www.nucleodoconhecimento.com.br/computer-engineering/biblical-citation

Artificial intelligence continues to demonstrate its innovative capacity in several scenarios: autonomous driving of vehicles; content recommendation; voice recognition; automatic translation; optimized planning and several other branches (NORVIG, 2013). At the same time, AI has shown robust results when applied to text analysis. (AGGARWAL, 2012) Therefore, a demonstration of this benefit that AI has brought in different areas is the recommendation systems that use artificial intelligence techniques and that such recommendation systems are tools adopted by successful business models, exemplifying: Amazon; Youtube and Google. (RICCI, 2015).

This article seeks to develop a hybrid recommendation system of biblical citation of the four gospels of the New Testament to subsidize the homiletics of religious heralds and the elaboration of sermons.

This article is divided into the following sections: Section 2.1 provides an overview of religious literature and artificial intelligence theory; Sections 2.2.1 and 2.2.2 describes related work in the academic literature Section 3 describes the methodology used; Section 4 presents the results obtained and the performance measurement of the recommendation system and Section 5 conclusions and future works.

2. THEORETICAL REVIEW

2.1 OVERVIEW OF BIBLICAL LITERATURE

Biblical literature is unique in history, consisting of 66 books in the Protestant version, it was written in three languages (Hebrew, Aramaic and Greek) and the period of making of the Holy Bible corresponds to about 1600 years. (COMFORT, 1998) (CHAPMAN, 2016)

Biblical literature is arranged in chapters (1,189) and verses (31,102)[8] in the Protestant version. It has several literary genres (narrative, poetic, wisdom, prophetic, biographical, apocalyptic, among others) (CHAPMAN, 2016) (ROGERSON, 2006). In addition to this diversity of literary genres, a multiplicity of linguistic phenomena is also

observed (metaphors, allegories, typology, among others). (ROBINSON e NOLIS, 2020).

Therefore, scholars and religious speakers choose to use resources for their understanding and explanation in a coherent way, for example hermeneutics. However, it is observed that in the academic literature techniques for automating biblical hermeneutics or homiletics are not yet applied, constituting a gap to be filled by works that provide this methodical study of the use of automation for understanding and communicating the biblical message.

2.2 OVERVIEW OF ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) is an extensive research field that is distributed in several subareas considering their particularities. The purpose of AI is to design intelligent machines or to solve problems that require logical or cognitive reasoning (NORVIG, 2013).

This article will be restricted to describing in an elementary way the areas of AI used to some degree in this proposal, namely: Data science, text mining, machine learning and recommendation systems.

- Data science: Data science is estimated as one of the most promising areas and the specialist in this field continues to be widely requested by the market (ROBINSON e NOLIS, 2020). Data science can be defined as “the art of using data to understand and solve real-world problems.” (ROBINSON e NOLIS, 2020). Therefore, the engine that powers data science is the data itself. Data are not knowledge in the strict sense of the word, that is, they contain knowledge, however, it is necessary that they be processed by some interventions so that they leave the “information” status and reach the “knowledge” status and for that, the KDD (Knowledge Discovery in Databases) which is included among the data science tools performs this task (CASTRO, 2017).

RC: 146578
Available in: https://www.nucleodoconhecimento.com.br/computer-engineering/biblical-citation

- Machine learning: Machine learning is defined as “the ability of computers to automatically improve their performance with experience, that is, learning through data” (MITCHELL, 2017). Machine learning is a solid mechanism of artificial intelligence and has some learning paradigms, with supervised and unsupervised learning being highlighted in this work. Supervised learning is based on trying to learn the mapping of inputs A to outputs B given a labeled set C, in other words, it is learning any function that predicts the output of new inputs, therefore, it requires the presence of a supervisor, or label, that guides the desired output in the learning process. The equation (1) summarizes this reasoning

$$C = \left(A_{i_1} B_i\right)_{i=1}^{N} \qquad (1)$$

where N is the number of examples.

In unsupervised learning, the objective is to find patterns in the data, that is, there is no figure of a supervisor, label or desired output, the organization of data in this paradigm is not known a priori, the equation (2) illustrates the idea of unsupervised learning.

$$A(x_i|\theta) \qquad (2)$$

Therefore, unsupervised learning is indicated in knowledge discovery and taxonomy problems. (NILSSON, 1998).

- Text mining: Text mining is analogous to data mining that seeks to find patterns in the data, however, respecting the specificities of textual data that are commonly unstructured in nature and with linguistic peculiarities. These singularities of textual data are mostly handled by Natural Language Processing

RC: 146578
Available in: https://www.nucleodoconhecimento.com.br/computer-engineering/biblical-citation

(NLP) that is contained in text mining. It is possible to note that conventional data mining does not meet the demand to scrutinize textual data, as they need a structured format for this, and it is worth noting that textual data correspond to 75% of the data produced by society (TANDEL, 2019). Consequently, text mining can play a leading role among data science tools.

Recommender System: Recommender systems (RS) arise with the need to perform a filter of the multiple possibilities of choice that a user has, since the world wide web (internet) has a wide spectrum of possibilities. Therefore, recommender systems are optimized search engine with customized items. The author MOHANTY (2020) conceptualizes recommender systems with the following sentence: “recommender systems are software tools and techniques that provide suggestions for items that are most likely of interest to a given user.” Recommender systems have some approaches, namely: collaborative filtering, content based and hybrid. Collaborative filtering is based on the idea that similar users share similar tastes or preferences, therefore, the emphasis here is on the user. The content-based approach seeks to relate and measure the similarity of the items’ content, in other words, similar items may be of interest to users who share the content of those items. The hybrid recommender system seeks to concatenate several perspectives of recommender system approaches, and the hybrid system is presented in the literature as a strategy to maximize the performance of recommender systems (MOHANTY, 2020).

This article will discuss in more detail the algorithms of each AI approach used in this work in the methodology section.

2.2.1 TEXT MINING APPLIED TO BIBLICAL LITERATURE

Some AI techniques are already applied to the biblical corpus, however, mostly for ancient or medieval writer recognition techniques, that is, these works aim to establish comparisons between paleography techniques, and use the biblical corpus due to its antiquity of authorship. (BRIA, 2019) (CILIA, 2020)

RC: 146578
Available in: https://www.nucleodoconhecimento.com.br/computer-engineering/biblical-citation

Other works seek to build translation machines and adopt the biblical text for this purpose (ESAN, 2020) (ASHENGO, 2021). However, very few works use the biblical corpus in an attempt to extract information from it (LUTZ, 2013) and when they do, they use more statistical representation techniques, (MURAI, 2013) rather than linguistic representation due to the complexity of the biblical corpus, falling, therefore, short of a minimally satisfactory interpretation.

It is possible to observe that there is a gap to be filled by academic works that use AI techniques applied to the biblical textual corpus, mainly in the area of meaning extraction or textual interpretation. Therefore, this article establishes this pioneering spirit in producing a recommendation system focused on the extraction of textual meaning and its understanding.

2.2.2 RECOMMENDATION SYSTEM APPLIED TO BIBLICAL LITERATURE

Recommender systems are widely used and necessary due to the substantial volume of information available in the digital world, so research is still needed to improve recommendation systems, as suggested by KHAN et al. (2020a). With the adoption of deep learning the limitations are concentrated in the lack of explainability of the models (black box) as assured by the authors (WANG, 2020). There are advances in some approaches to recommendation systems, for example: collaborative filtering in the work of (SHICHANG, 2021) had a significant improvement in its performance and optimization for news recommendations, using hybrid modeling based on user items.

Several authors declare that the recommendation system that combines different methodologies, that is, the hybrid systems, is the approach that converges to the best performance of (KHAN et al., 2020a) recommendations. The work by (RAMESH, 2022) follows the same conclusion and points out that hybrid systems, the combination of "collaborative filtering" and "content filtering", enhance the effectiveness of a recommendation system, therefore, they proposed the development of a hybrid model to make this finding. Recommender systems for texts or literatures remain, in

comparison with other applications, a small and developing field. Some areas in which they are applied are: education (ZHU et al., 2021) (DIAS, 2020) (ZHU, 2021) (CHANANA, 2019); law (LIU, 2019) (OSTENDORFF et al., 2021) (SAXENA, 2021) (THOMAS et al., 2020) (DHANANI, 2022); social networks (RENJITH, 2021) (IRFAN et al., 2019); news (ASHENGO, 2021); e-commerce (KHAN et al., 2020b) (SHRIVASTAVA, 2019) and among other rudimentary applications. This work will focus on the educational perspective.

Recommender systems for educational purposes aim to “research and select relevant content” (PINHO et al., 2019). According to PINHO et al. (2019), the target audience of educational recommendation systems is students, around 44.44% and 16.67% are teachers. The most used techniques were: profile-based collaborative filtering, context-based filtering, content-based filtering and applied on a smaller scale to hybrid filtering and the authors PONTE et al. (2021) conclude in a similar way, however, with the addition that the educational content recommendation system collaborates with learning and and both works highlight a lag in the use of semantic extraction techniques , in the articles analyzed only 1% used semantics. The authors PINHO et al. (2019) also propose for future work the use of educational recommendations for books and articles from university libraries. Regarding gaps and future work, the authors emphasized the adoption of more hybrid systems and the adoption of ontologies. The authors LIU et al. (2022) use deep learning for their educational recommendation software, with the techniques applied: CNN; RNN; LSTM etc. They emphasize that the limitation of the use of deep learning for recommender systems are the unexplainable decisions (black box), the authors assure that there is room for many contributions in recommendation systems for educational purposes, since the field is relatively new and there are several gaps. For the authors KHANAL et al. (2020) the methodology that presents the best competitive advantage for an educational recommendation system is the hybrid approach.

3. METHODOLOGY

RC: 146578
Available in: https://www.nucleodoconhecimento.com.br/computer-engineering/biblical-citation

The present research uses the hybrid approach for the construction of its biblical citation recommendation system, which is the association of perspectives: based on content and collaborative filtering.

3.1 CONTENT-BASED APPROACH

Within the content-based approach, a rigorous analysis of the linguistic aspects of the biblical corpus is carried out, unlike what is most often done by authors in academia, that is, a statistical representation of textual data through methods such as BOW or TF-IDF, (WEISS, 2005) in the respective research, a semantic representation of the text was chosen through the topic modeling algorithm LSI (Latent Semantic Indexing). LSI uses the structure of the SVD matrix factorization system (Singular Value Decomposition) whose equation is shown below.

$$X = U\sum V^T \qquad (3)$$

Where U is an m matrix, V is an n matrix and is a diagonal matrix. With the use of LSI, the problem of synonyms and polysemy that are observed in writings or textual data is addressed (AGGARWAL; ZHAI, 2012).

The LSI forms the basis of the PLN (Natural Language Processing) used in this proposal and consequently enables the extraction of textual content from the biblical corpus. To implement the textual content similarity search engine, the cosine similarity measure was adopted because it is the most used for this purpose in the academic literature (AGGARWAL, 2012). Equation (4) exposes the cosine similarity measure.

$$cos(U,V) = \frac{\sum_{i=1}^{k} f(u_i) \cdot f(v_i)}{\sqrt{\sum_{i=1}^{k} f(u_i)^2} \cdot \sqrt{\sum_{i=1}^{k} f(v_i)^2}} \quad (4)$$

Where: ui and vi are components of the vector U and V respectively.

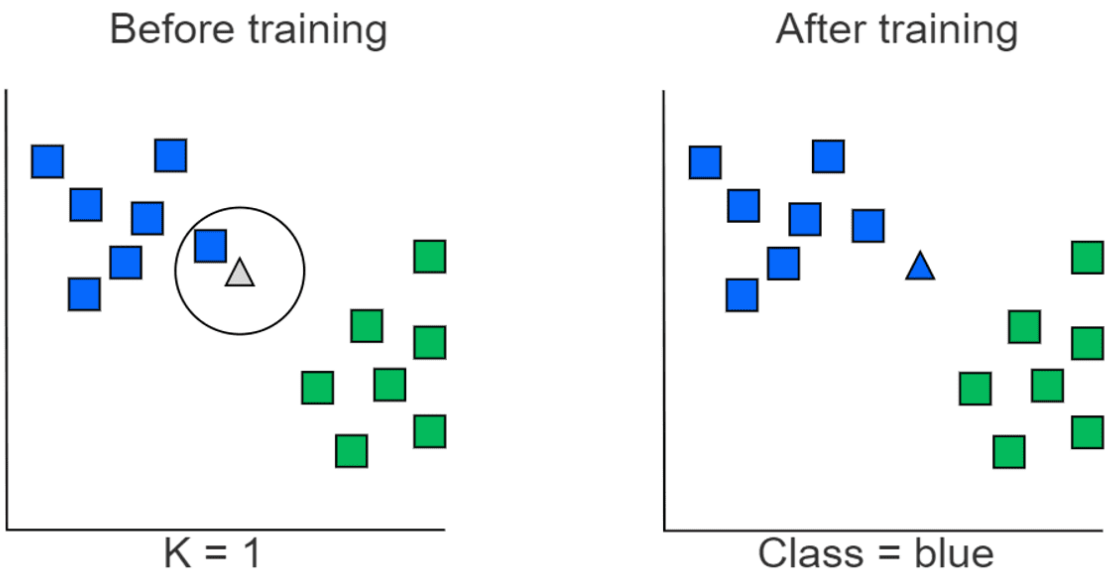
This content-based approach was implemented in the Python language, using the classic natural language processing libraries *Spacy* and *Gensim*.

3.2 COLLABORATIVE FILTERING APPROACH

In the collaborative filtering approach, three algorithms were used for performance comparison, namely KNN (K-Nearest Neighbors), SVD (Singular Value Decomposition) and Baseline. They work by performing regression in the supervised learning paradigm, that is, bringing a numerical output. Despise KNN being a classification algorithm, regression can be performed with it. Its way of working derives from choosing a K number and measuring the distance between the new data to be labeled and the data contained in the perimeter of the K number. Figure 1 demonstrates this, and equation (5) summarizes the Euclidean distance measure.

Figure 1 – KNN Algorithm

RC: 146578
Available in: https://www.nucleodoconhecimento.com.br/computer-engineering/biblical-citation



Source: Authors.

$$d = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} \qquad (5)$$

Where: n is the size of the samples. The KNN was adopted to measure the similarity by distance of the scores of the texts.

The SVD used in this collaborative filtering approach is analogous to the SVD of the content-based approach, using equation (3) As the collaborative filtering matrices have a sparse characteristic, the use of SVD acts in the reduction of their dimensionality (AGGARWAL, 2012).

The Baseline algorithm is a basic regression algorithm, and its choice was due to the fact that the authors sought to compare some types of algorithms and analyze their performance in the uniqueness of the biblical text.

All algorithms used in this collaborative filtering approach were implemented in the Python programming language using the library *Surprise*.

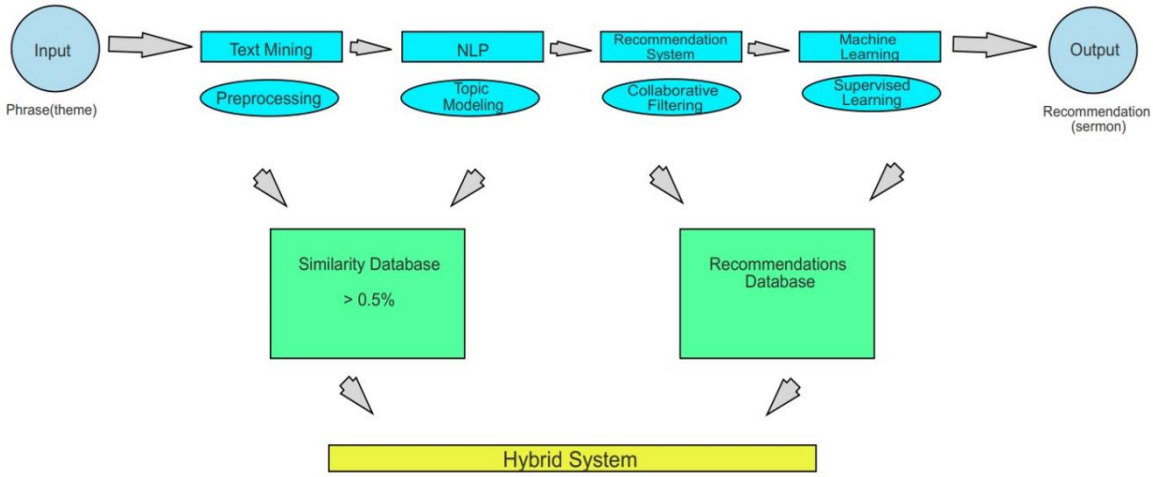
3.3 TRAINING

The training phase is divided into two parts:

- Linguistic information extraction: in this step, the entries (thematic search) are inserted in the LSI in parallel with the biblical corpus and consequently a search is made by textual similarity (entry versus biblical corpus). The similarity cut-off level is 50%, that is, biblical texts that have a similarity greater than 50% to the search terms. The texts whose similarities reach the established target feed a preference or ranking table.
- Collaborative filtering: in this training phase, the similarity ranking table is provided to the algorithms: KNN, SVD and Baseline and they process preferences for citation recommendations. For KNN the number of k = 25 was chosen, SVD was trained in 35 epochs and Baseline was trained in 10 epochs. Figures 2 and 3 illustrates this process.

Figure 2 – Hybrid recommender system training overview

RC: 146578
Available in: https://www.nucleodoconhecimento.com.br/computer-engineering/biblical-citation



Source: Authors.

3.4 RESULTS EVALUATION METRICS

For problems of this nature, conventional metrics like RMSE (Root Mean Squared Error), equation (6), and MAE (Mean Absolute Error), equation (7), were used to evaluate regression problems (Mohanty et al., 2020)

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(PredictedActual)^2}{N}} \quad (6)$$

$$MAE = \frac{1}{N}\Sigma|PredictedActual| \quad (7)$$

Data were divided between training and testing, and cross-validation was adopted in 10 folders to avoid overfitting the model.

RC: 146578
Available in: https://www.nucleodoconhecimento.com.br/computer-engineering/biblical-citation

3.5 DATABASE

The database used in this article is the four gospels of the biblical New Testament

(Gospel of Matthew, Gospel of Mark, Gospel of Luke and Gospel of John). The version adopted is the Protestant King James Bible Version in English, the version most used in related works.

The data mining treatment employed was the removal of stop words, unigrams and Removing words that appear only once

4. RESULTS

The results obtained converge with the objective of this article, which was the proposition of biblical quotations for the construction of religious sermons and, through this, provide, in an automated way, subsidy for the communication and teaching of the religious leader. Table 1 presents these results.

Table – Biblical quote recommendations

Algorithm	Search term	Recommended Bible passages	Theme of recommended texts
SVD	Eternal life	Lc 1; Mt 10; Mt 1; Lc 12; Mt 6; Mt 5; Lc 17; Mt 23; Lc 21; Lc 6	Bodily adultery.
	Loving God and neighbor	Mc 16; Mt 12; Jo 16; Mt 18; Lc 6; Lc 11; Lc 13; Lc 17; Mt 11; Lc 12	Left out of the kingdom of God.
	Blasphemy against God	Jo 19; Jo 1; Lc 24; Mc 1; Jo 9; Mt 27; Mt 9; Mc 2; Mc 5; Jo 18	Eternal damnation.
	Hypocrisy of the Pharisees	Mt 18; Mc 16; Mt 12; Lc 11; Jo 16; Lc 13; Lc 6; Mt 11; Mt 16; Mt 13.	Prayers and fasts and good works.
KNN	Eternal life	Jo 2; Mt 14; Jo 20; Mc 15;	Repentance of transgressions.

RC: 146578
Available in: https://www.nucleodoconhecimento.com.br/computer-engineering/biblical-citation

		Jo 18; Mt 27; Jo 9 Lc 24; Jo 19; Jo 1	
	Loving God and neighbor	Lc 23; Lc 5; Mt 8; Mc 7 Mc 6; Lc 16; Lc 15; Lc 8; Lc 7; Mc 3	Blood of the New Covenant.
	Blasphemy against God	Jo 2; Mt 14; Jo 20; Mc 15; Jo 18; Mt 27; Jo 9; Lc 24; Jo 19; Jo 1	Repentance of transgressions.
	Hypocrisy of the Pharisees	Jo 16; Lc 6; Mt 12; Mc 16; Lc 17; Mt 18; Lc 11; Lc 12; Lc 13; Lc 1	Beatitudes of the Kingdom of God.
BASELINE	Eternal life	Jo 2; Mt 14; Jo 20; Mc 15; Jo 18; Mt 27; Jo 9 Lc 24; Jo 19; Jo 1	Repentance of transgressions.
	Loving God and neighbor	Lc 21; Mt 23; Mt 5; Mc 13; Mt 6; Mt 24; Mt 1; Mt 10; Lc 1; Lc 12	God's kingdom.
	Blasphemy against God	Lc 18; Mc 9; Lc 14; Jo 15; Mc 4; Mt 25; Mt 19; Mc 10; Lc 3; Mt 20	They loved the glory of men more than the glory of God.
	Hypocrisy of the Pharisees	Lc 23; Lc 5; Mt 8; Mc 7; Mc 6; Lc 16; Lc 15; Lc 8; Lc 7; Mc 3	Blood of the New Covenant.

Source: Authors.

Religious sermons usually follow a thematic contextual flow and the results achieved by the recommendation system proposed in this article show this ability. The thematic search "Eternal life" returned the following citation recommendations: "Avoid losing eternal life", "Characteristics of saved" and "Actions of the saved" and this set of citations recommended by the implemented system brings a context knowledge,

because through a macro theological theme is extracted from micro theological themes linked to the macro theme. Table 2 demonstrates this finding.

Table 2 – Contextual scope of recommendations

Algorithm	Search term	Recommended Bible passages	Theme of recommended texts
	Theological concept macro (Search term)	Theological concept micro (Recommendations - returned terms)	Contextual perimeter
SVD / KNN	Eternal life	I. Avoid losing eternal life II. Characteristics of saved III. Actions of the saved	Holistic /remote
BASELINE	Repentance	I. Repentance II. Repentance III. Repentance	Restricted /immediate

Source: Authors.

Two levels of contextual understanding are observed by the recommendations of the system built in this article, that is, remote context and immediate context. The SVD/KNN algorithms demonstrated a more holistic behavior (remote) of contextual understanding and the Baseline algorithm presented a more restricted contextual understanding (immediate).

The metrics obtained demonstrate a low level of error in predictions, corroborating the robustness of the implemented recommendation system. The table 3 presents the metrics.

Table 3 – Performance metrics – Average of k folders (k=10)

	RMSE	MAE
SVD	0.931	0.776
KNN	0.938	0.734
BASELINE	0.968	0.783

Source: Authors.

RC: 146578
Available in: https://www.nucleodoconhecimento.com.br/computer-engineering/biblical-citation

The Baseline algorithm presented similar performance (MAE 0.783) and with a lower training cost than the SVD algorithm (MAE 0.776).

However, the Baseline algorithm brings a more restricted contextual result to the detriment of other more robust algorithms (SVD/KNN) that converge to broader results in terms of perimeter. Another point to be considered in relation to the Baseline algorithm is that the New Testament gospels present a predominantly narrative literary genre, therefore, this baseline behavior can be different in biblical texts of more diverse genres.

These findings demonstrate that with advanced techniques of natural language processing combined with machine learning techniques it is possible to extract or retrieve semantic contextual information from widely complex textual corpus as is the case of the biblical corpus. However, it is necessary to emphasize that these results were achieved with this combination of methodologies, that is, using a hybrid methodology.

5. CONCLUSION

Religion gained a relevant prominence in the life of ancient and modern society, and the Holy Bible is consecrated the most significant literary work of Christianity, therefore, its message needs to be communicated in a satisfactory way, and religious leaders need to have tools that enhance the communication of this message.

The biblical citation recommendation system proposed in this article aims to deliver an automated and optimized way of recommending biblical citations with the use of artificial intelligence and natural language processing despite all the complexity contained in the biblical corpus and thereby subsidizing the construction of biblical sermons and religious homiletics.

Therefore, the hybridization of NLP techniques and machine learning through the platform of a recommender system can favor the recovery of contextual information from textual corpus of high semantic complexity similarly to biblical literature.

The results achieved in this research showed a capacity for contextual discernment, based on the following facts:

- The recommendations follow a semantic understanding of the theological terms searched for the reason that the recommended biblical passages show a holistic understanding of the searched theme.
- The recommendations have an inclination towards symmetry and contextual harmony with the researched macro theological concepts.
- The RMSE and MAE metrics showed good performance and low error in predictions, that is, values below 1. The cross-validation carried out demonstrates that there was no overfitting in the test data.

Future works: for the advancement of the research, the authors propose the following steps:

- Perform more comprehensive natural language processing, involving entity recognition, entity relationships, POS, among other NLP techniques. Aiming to improve the extraction of linguistic minutiae from the biblical corpus.
- Investigate the possibility of using embeds trained with the biblical corpus.
- Investigate the possibility of expanding the hybridization of the recommendation system.

ACKNOWLEDGMENT

This article was developed with the support of Capes and Mackpesquisa.

BIBLIOGRAPHY

AGGARWAL, C. Charu.; ZHAI, Cheng. X. **Mining Text Data**. 1st. ed. Springer, 2012.

ASHENGO, Yeabsira A.; AGA, Rosa T.; ABEBE, Surafel. L. Context based machine translation with recurrent neural network for english–amharic translation. **Machine Translation**, v. 35, n. 19-36, 2021.

BRIA, Alessandro.; CILIA, Nicole, D.; STEFANO, Claudio.; FONTANELLA, Francesco.; MARROCOS, Claudio.; MOLINARA, Mario.; FRECA, Alessandra.;

TORTORELLA, Francesco. Deep transfer learning for writer identification in medieval books. **2018 Metrology for Archaeology and Cultural Heritage (MetroArchaeo),** n. 455-460, 2019.

CASTRO, Leandro. N. de; FERRARI, D. G. **An Introduction to Data Mining: Basic Concepts, Algorithms and Applications**. 1rd. ed.: Saraiva Educação SA, 2017.

CHANANA, Rishabh.; SINGH, Amit K.; VERMA, Ankita. Recommending Relevant Research Articles based on Citation Graph and Textual Information. **5th IEEE International Conference on Signal Processing, Computing and Control (ISPCC 2k19),** Oct 10-12, 2019, JUIT, Solan, India, 2019.

CHAPMAN, Stephen. B.; SWEENEY, Marvin. A. **The hebrew bible/old testament**. 1st. ed. Cambridge University Press, 2016.

CILIA, Nicole D.; STEPHANO, Claudio.; FONTANELLA, Francisco.; MARROCCO, Claudio.; MOLINARA, Mario.; FRECA, Alessandra S. An experimental comparison between deep learning and classical machine learning approaches for writer identification in medieval documents. **Journal of Imaging**, v. 6, n. 89, 2020.

COMFORT, P. W. **A Origem e Autenticidade da Bíblia**. 1rd. ed. Casa Publicadora das Assembleias de Deus (CPAD), 1998.

DHANANI, J.; MEHTA, R.; RANA, D. Effective and scalable legal judgment recommendation using pre-learned word embedding. **Complex Intelligent Systems (2022)**, 2022.

DIAS, Laura.; BARRERE, Eduardo.; SOUZA, Jairo. The impact of semantic annotation techniques on content-based video lecture recommendation. **Journal of Information Science**, 2020.

ESAN, Adebimpe.; OLADOSU, John.; OYELEYE, Christopher, ADEYANJU, Ibrahim.; OLANIYAN, Olatayo, OKOMBA, Nnamdi.; OMUDUNBI, Bolaji.; et al. Development of a recurrent neural network model for english to yoruba machine translation. **International Journal of Advanced Computer Science and Applications**, v. 11, n. 602-609, 2020.

IRFAN, Rizwana. et al. **SocialRec:** A Context-Aware Recommendation Framework With Explicit Sentiment Analysis. **IEEE Access**, 2019.

ATEN, Jaime D.; SMITH, Wendy, R.; DAVIS, Edward.; et al. The psychological study of religion and spirituality in a disaster context: A systematic review. **Psychological Trauma**, v. 11, n. 597–613, 2019.

KHAN, A. et al. Systematic Review on Recommendation Systems. **2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN),** 2020.

KHAN, Z. et al. Contextual recommender system for E-commerce applications**. Applied Soft Computing**, 2020.

KHANAL, S. et al. A systematic review: machine learning based recommendation systems for e-learning. **Education and Information Technologies**, 2020.

LIU, T. et al. A review of deep learning-based recommender system in e-learning environments**. Artificial Intelligence Review**, 2022.

LIU, Y.; LUO, X.; YANG, X. Semantics and Structure Based Recommendation of Similar Legal Cases. **2019 IEEE 14th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)**, 2019.

LUTZ, D. S. The relative influence of european writers on late eighteenth-century american political thought. **The American Political Science Review**, v. 78, n. 189–197, 1984.

LUTZ, D. S. Exegetical science for the interpretation of the bible: Algorithms and software for quantitative analysis of christian documents. **Studies in Computational Intelligence**, v. 492, n. 67-86, 2013.

MITCHELL, T. M. **Introduction To Machine learning An early draft of a proposed Textbook**. 1rd. ed. [S.l.]: McGraw Hill Education, 2017.

MITKOV, R. **The Oxford Handbook of Computational Linguistics**. 1st. ed. Oxford Press, 2003.

MOHANTY, Sachi.; CHATTERJEE, Jyotir.; JAIN, Sarika.; ELNGAR, Ahmed.; GRUPTA, Priya. **Recommender System with Machine Learning and Artificial Intelligence Practical Tools and Applications in Medical, Agricultural and Other Industries**. 1rd. ed. Wiley, 2020.

MURAI, H. Exegetical science for the interpretation of the bible: Algorithms and software for quantitative analysis of christian documents. **Studies in Computational Intelligence** 492:67–86, (2013)

NILSSON, N. J. **Artificial Intelligence A New Synthesis**. 1st. ed. Morgan Kaufmann Publishers, 1998.

NORVIG, P. **Artificial Intelligence**. 1rd. ed. Publishing company: Grupo GEN, 2013.

OSTENDORFF, M. et al. Evaluating Document Representations for Content-based Legal Literature Recommendations. **ICAIL '21: Proceedings of the Eighteenth International Conference on Artificial Intelligence and Law**, 2021.

PINHO, P. et al. Developments in Educational Recommendation Systems: a systematic review. **2019 IEEE Frontiers in Education Conference (FIE)**, 2019.

PONTE, M.; ZORRILLA, A.; RUIZ, I. Recommendation Systems for Education: Systematic Review. **Electronics**, 2021.

RAMESH, R.; VIJAYALAKSHMI, S. Improvement to Recommendation system using Hybrid techniques. **2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)**, 2022.

RENJITH, S.; SREEKUMAR, A.; JATHAVEDAN, M. SemRec—An efficient ensemble recommender with sentiment based clustering for social media text corpus. **Concurrency Computat Pract Exper.**, 2021.

RICCI, Francisco.; ROKACH, Lior.; SHAPIRA, Bracha. KANTOR, Paul. **Recommender Systems Handbook**. 1rd. ed. Springer, 2015.

ROBINSON, E.; NOLIS, J. **Build a Career in Data Science**. 1st. ed. Manning Publications Co., 2020.

ROGERSON, J. W.; LIEU, J. M. **The Oxford Handbook of Biblical Studies**. 1rd. ed. Oxford University Press, 2006.

SAXENA, K.; SUNKLE, S.; KULKARNI, V. Towards Recommendations from User-specific Insights based on Historical Legal Cases. **14th Innovations in Software Engineering Conference**, 2021.

SEN, L. C. H. E.; BROWNE, D. T. Keeping the faith: Religion, positive coping, and mental health of caregivers during covid-19. **Frontiers in Psychology**, v. 12, n. 67–86, 2022.

SHICHANG, Z. Research on Recommendation Algorithm Based on Collaborative Filtering. **Research on Recommendation Algorithm Based on Collaborative Filtering**, 2021.

SHRIVASTAVA, R.; SISODIA, D. Product Recommendations Using Textual Similarity Based Learning Models. **2019 International Conference on Computer Communication and Informatics (ICCCI -2019), Jan. 23 – 25, 2019, Coimbatore, INDIA**, 2019.

TANDEL, S.; JAMADAR, A.; DUDUGU, S. A survey on text mining techniques. **International Conference on Advanced Computing Communication Systems (ICACCS2019)**, 2019.

THOMAS, M. et al. Quick Check: A Legal Research Recommendation System. **NLLP @KDD 2020**, 2020.

WANG, M.; LIU, X.; JING, L. Deep Learning based Recommedation System: A Review of Recent Works. **2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC 2020)**, 2020.

WEISS, S.; INDURKHYA, N.; ZHAN, T.; DAMERAU, F. J. **Text Mining Predictive Methods for Analyzing Unstructured Information**. 1rd. ed. Springer, 2005.

ZHU, J.; PATRA, B.; YASEEN, A. Recommender system of scholarly papers using public datasets. **AMIA Jt Summits Transl Sci Proc.**, 2021.

ZHU, Y. et al. Recommending scientific paper via heterogeneous knowledge embedding based attentive recurrent neural networks. **Knowledge-Based Systems**, 2021.

APPENDIX - FOOTNOTE

6. https://www.cia.gov/the-world-factbook/field/religions/.

7. https://www.guinnessworldrecords.com/world-records/best-selling-book-of-non-fiction.

8. https://www.biblebelievers.com/believers-org/kjv-stats.html.

Sent: June 1, 2023.

Approved: June 22, 2023.

---

[1] Graduated in Computer Network Technology from Universidade Nove de Julho (2013-2015). Master in Electrical Engineering from Universidade Presbiteriana Mackenzie (Mackgraphe) and PhD in Electrical Engineering from the same (LCoN). He has experience in the area of Computer Science, with an emphasis on Computer Networks and Artificial Intelligence. ORCID: https://orcid.org/0000-0002-4812-0963.

[2] Full Professor at Universidade Presbiteriana Mackenzie, holds a degree in Mechanical Engineering from the Technological Institute of Aeronautics (1974), a Master's degree in Applied Mathematics - Informatics from the Technological Institute of Aeronautics (1979) and a PhD in Informatics from the Pontifical Catholic University of Rio de Janeiro (1989). Full Professor at ITA (retired). ORCID: https://orcid.org/0000-0001-8361-2165. Currículo Lattes: http://lattes.cnpq.br/2067336430076971.

[3] He is currently a researcher at Universidade Presbiteriana Mackenzie. He has experience in the area of Electrical Engineering, with an emphasis on Superconducting Materials and Devices. ORCID: https://orcid.org/0000-0001-5798-5523. Currículo Lattes: http://lattes.cnpq.br/2387328257612150.

[4] Electrical Engineer from the Federal University of Goiás (1996), Master in Electrical Engineering from Unicamp (1998), PhD in Computer Engineering (2001) from Unicamp, MBA in Strategic Business Management from the Catholic University of Santos (2008) and made Empretec by Sebrae-SP (2012). Professor and researcher at the Master's Program in Informatics at Unisantos, and at the Graduate Program in Electrical and Computing Engineering at Mackenzie. ORCID: https://orcid.org/0000-0003-3409-4589. Currículo Lattes: http://lattes.cnpq.br/2741458816539568.

[5] Advisor. Graduated in Mathematics-Informatics from UFJF - Federal University of Juiz de Fora (1994), Master in Science - Informatics from ITA - Technological Institute of Aeronautics (1997) and PhD in Electrical Engineering from POLI-USP - Polytechnic School of the University of São Paulo Paul

(2003). ORCID: https://orcid.org/0000-0001-8029-072X. Currículo Lattes: http://lattes.cnpq.br/3894359521286830.