AUTOMATIC SUMMARIZATION 2

Project 29

ROSHAN FERNANDO
KAVINDU RAVISHAN
ALI GOODARZI

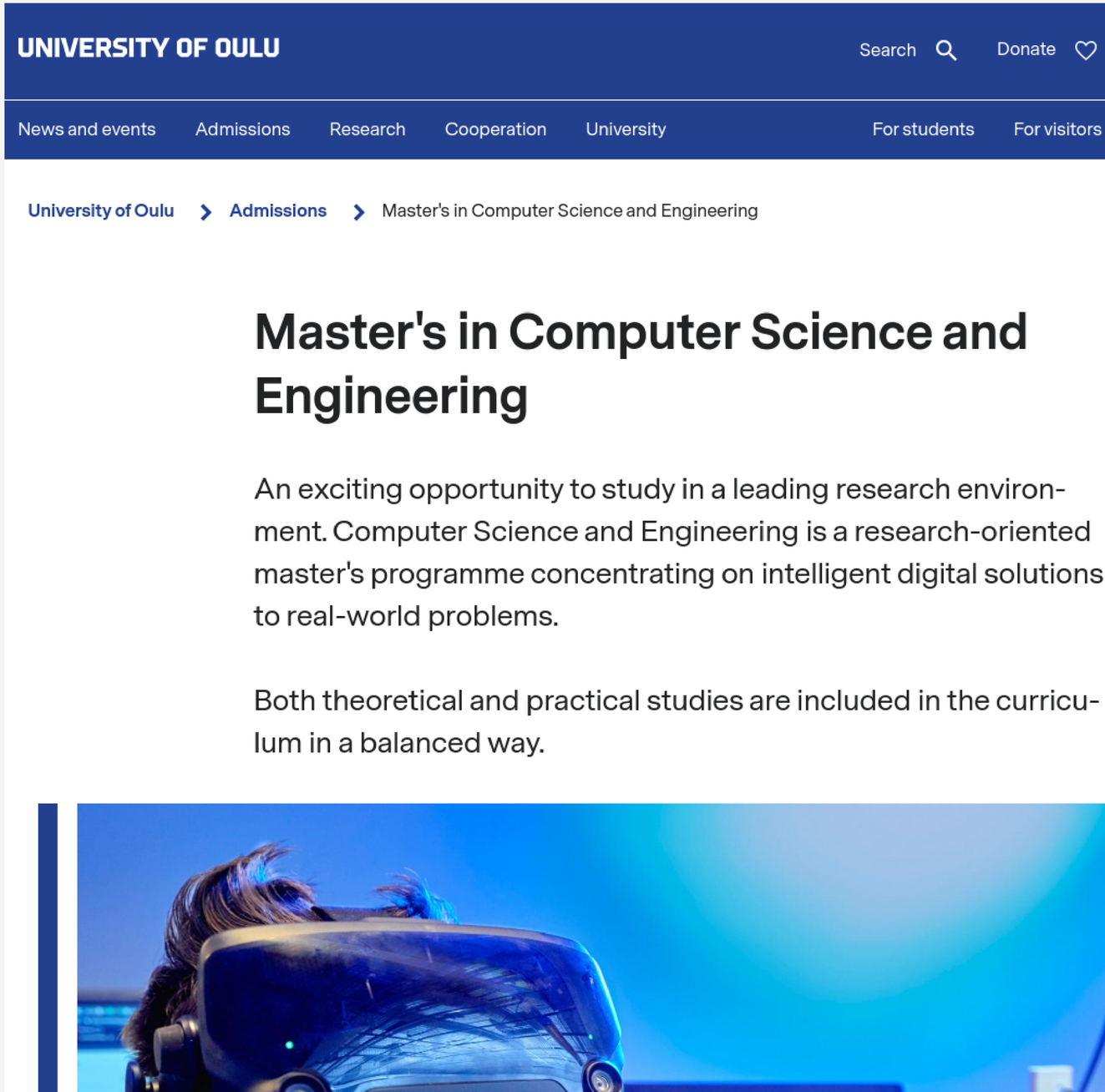
TECHNOLOGIES

OBJECTIVES

- A Python-based system for automatic text summarization.
- Extract keywords from HTML documents.
- Analyze word frequency and identify named entities in a PDF document (Research Paper).
- Summarize documents considering frequent wording (minus stopwords) and named-entities.
- Summarize documents and evaluate using ROUGE metrics.
- Improve summarization by minimizing redundancy among sentences.
- Shift from histogram frequency to RAKE for keyword choice.
- Test multiple summarization algorithms (sumy) on the same document.
- Evaluate the performance of various summarization techniques on a set of scientific papers from an Elsevier journal.

TASK 01

EXTRACTING KEY INFORMATION FROM WEB PAGE



Here we extracted key information from an web page. It takes a URL as input and returns a dictionary with three keys: Title, Abstract & Subsection Titles.

https://www.oulu.fi/en/apply/masters-computer-science-and-engineering

Methodology

- Use requests to fetch HTML content.
- Parse with BeautifulSoup.
- Extract data from tags: <title>, <p class='abstract'>, <h2>, <h3>, <h4>.

Data Cleaning

- Remove extra spaces.
- Format subsections as bullet points.

	Extracted Content
Title	Master's in Computer Science and Engineering | University of Oulu
Abstract	Top reasons to study Computer Science and Engineering Choose your orientation or mix all four: Arti
Subsection Titles	• Degree title • Study places • Duration of studies • Scope • Teaching method • Next application period • Top reasons to study Computer Science and Engineering • Apply at Studyinfo to study Computer Science and Engineering in Oulu • Artificial intelligence research • Core skills and competence • Study Applied Computing • Study Artificial Intelligence • Study Computer Engineering • Study Cyber Security • Programme structure and courses • Admissions criteria • Admissions Criteria for Master's in Computer Science and Engineering • Occupational profiles of the graduates • A new journey with the CSE master’s programme at the University of Oulu • Oulu: The Silicon Valley of Northern Scandinavia? • Video - Study Computer Science and Engineering in Oulu • A solid foundation for the future careers • Apply online to study Computer Science and Engineering • Applying to Master's Programmes • International Programmes • Chat with us

TASK 02

ANALYZING TEXT FROM A PDF: WORD FREQUENCY HISTOGRAM

Expert Systems With Applications 238 (2024) 122016

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Towards job screening and personality traits estimation from video transcriptions

Yazid Bounab [a], Mourad Oussalah [a,*], Nabil Arhab [a], Salah Bekhouche [b]

[a] University of Oulu, Faculty of ITEE, CMVS, Finland
[b] University of the Basque Country, UPV/EHU, Spain

ARTICLE INFO

Keywords:
Big-five personality
Deep learning
Human behavior

ABSTRACT

In recent years, natural language processing (NLP) has gained new territory beyond its traditional use in text mining applications. This paper shows the effectiveness of NLP techniques in assessing the apparent human personality from his/her video transcript, building a bridge between NLP and computer vision-based reasoning.

In this paper, a new deep learning model using attention mechanism and bidirectional LSTM layers for estimating the Big-five personality traits is provided and then tested on ChatLearn video dataset. The robustness of the approach is then tested by generalizing the method to two other datasets (B5 corpus and Mypersonality datasets). Several empirical evaluations taking into account the various inputs of the data processing pipeline have been performed to yield optimal model parameters.

The developed model is tested on the APA'2016 competition dataset from Chalearn V2 Challenge Workshop for both Big-five personality trait estimation and interview score estimation. We achieved an average of 89.27% in personality trait recognition rate and 89.16% score in the Job Interview challenge. Similar trends of high accuracy score estimation is held for B5-corpus and MyPersonality datasets as well.

The results outperformed several state-of-art approaches, demonstrating our approach's feasibility in extending the computer vision approach of first impression personality to natural language processing. This opens up a new direction in multimedia analysis.

1. Introduction

In the last two decades there has been a surge of interest in affective computing and personality computing that seek to automatically recognize and synthesize individual personality (Cowie et al., 2001). This has been performed either through a multi-modal combination of facial images, speaking style, body movement, writing style, or a single modality of the above (Vernon, Sutherland, Young, & Hartley, 2014). In this context, the personality calculus primarily focused on estimating human personality through different factors yielding different psychological models such as the Big-Five personality traits and Myers-Briggs Type Indicator models. The Big-Five personality traits, often referred to as the Five Factors Model (or the Big-Five) (Roberts, Kuncel, Shiner, Caspi, & Goldberg, 2007): Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness, is among the most investigated models in affective computing literature. Especially, the Big-Five model is found to prognosticate many life aspects such as work performance, interpersonal relations, emigration, social beliefs, and well-being (Cherry, 2019). We distinguish two research streams in estimating individual personality prediction (Matthews, Deary, & Whiteman, 2009). The first one advocates a correct recognition of essential personality traits using self or acquaintance reports that often involves interviews and/or successive observations. The second stream boils down to the process of automatic recognition of the personality traits of an unfamiliar individual using computational methods. Typically, computational psychology research primarily deals with this second stream as it attempts to estimate personality traits from videos and multimedia clips highlighting an individual's behavior.

In this context, one may mention the INTERSPEECH 2012 Speaker Trait Challenge (Schuller et al., 2012), which provides an audio dataset and a set of extracted features to enable comparison of computational methods for the estimation of Big-Five personality traits. With the proliferation of social media platforms, a growing interest has been noticed in predicting apparent personality traits from social media content. Biel, Teijeiro-Mosquera, and Gatica-Perez (2012) used YouTube vloggers video frames to estimate personality impressions using facial expressions. Likewise, Cristani, Vinciarelli, Segalin, and Perina (2013)

* Corresponding author.
E-mail addresses: Yazid.Bounab@oulu.fi (Y. Bounab), Mourad.Oussalah@oulu.fi (M. Oussalah), Nabil.Arhab@oulu.fi (N. Arhab), bekhouchesalah@gmail.com (S. Bekhouche).

https://doi.org/10.1016/j.eswa.2023.122016
Received 4 May 2023; Received in revised form 28 September 2023; Accepted 2 October 2023

Key Features:

- Extract text from PDF files.
- Generate word frequency histogram, excluding stopwords.
- Identify person and organization named entities.

Methodology

- Use spacy for natural language processing.
- en_core_web_sm model for English language processing.
- Filter out stopwords, punctuation, and whitespace.
- Display top 20 most frequent words as a histogram.

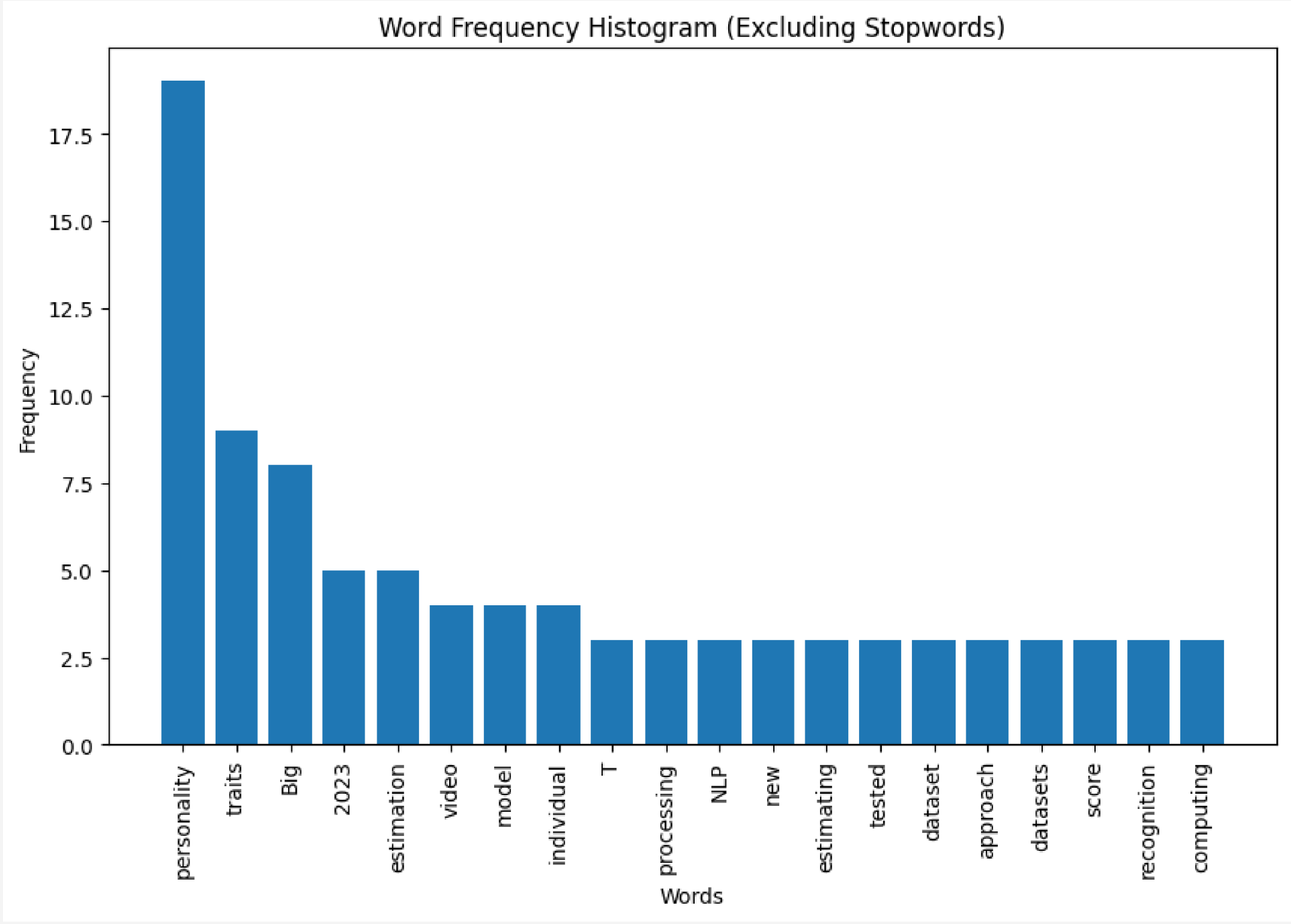
Visualization

- Plot histogram using matplotlib.
- Display word frequencies and named entities in tables using PrettyTable.
- Execution: Analyze /content/task-02/paper.pdf.

Word	Frequency
personality	19
traits	9
Big	8
2023	5
estimation	5
video	4
model	4
individual	4
T	3
processing	3
NLP	3
new	3
estimating	3
tested	3
dataset	3
approach	3
datasets	3
score	3
recognition	3
computing	3

spaCy

Persons	Organizations
Yazid Bounaba	Elsevier Ltd.
Mourad	ScienceDirect
Nabil Arhaba	CMVS
Salah Bekhoucheb	bUniversity
Roberts	the Basque Country
Caspi	NLP
Yazid.Bounab@oulu.fi	NLP
Y. Bounab	NLP
Mourad.Oussalah@oulu.fi	ChatLearn
Nabil.Arhab@oulu.fi	Chalearn V2 Challenge Workshop
N. Arhab	MyPersonality
(S. Bekhouche).in	Vernon
Teijeiro-Mosquera	Myers
Cristani	Kuncel
Vinciarelli	Conscientiousness
https://doi.org/10.1016/j.eswa.2023.122016	Openness
Received	
	Matthews, Deary, &
	Whiteman
	Trait Challenge
	Schuller
	Gatica-Perez
	YouTube



TASK 03

DOCUMENT SUMMARIZATION USING TF-IDF AND POS WEIGHTS IN SPACY

Features

- Sentence-level analysis treating each sentence as a sub-document.
- Compute TF-IDF scores for words and named entities.
- Calculate weighted sentence scores using word, named-entity TF-IDF, and positional weights.

Methodology

- Use spacy for NLP tasks: tokenization, lemmatization, named-entity recognition.
- en_core_web_md model for English processing.
- Word-level TF-IDF via TfidfVectorizer.

Sentence weight formula

$$
\begin{aligned}
&\text{Sweight} \\
&= \\
&(\sum \text{word TF-IDF}) \\
&+ \\
&2(\sum \text{named-entity TF-IDF}) \\
&+ \\
&\text{POS weight.}
\end{aligned}
$$

Sentence	Weight
article address follow research question RQ1 study major benefit blockchain context industry 4.0 RQ2 identify study major driver enabler Blockchain technology industry 4.0 RQ3 study associate blockchain capability successful industry 4.0 implementation perspective RQ4 study different industry 4.0 sphere sub domain Blockchain technology realisation rq5 identify study major application Blockchain technology industry 4.0	9.29
datum analytic Artificial Intelligence AI understand case application	5.69
process take information previously store Enterprise Resource Planning ERP company	5.31
paper aim study significant potential role Blockchain Industry 4.0	5.15
instance Leng et al	4.96
need study identify potential role Blockchain Industry 4.0	4.92
supply chain dynamic structure comprise multiple business work satisfy customer need add value raw material level final product	4.22
Blockchain Blockchain define decentralised distribute directory drive smart contract provide opportunity traceability aid record management automation supply chain payment application business transaction	4.12
main chain support secondary tertiary chain co operate form ecosystem supply chain network 11,12	4.04
provide inclusive distribution strategy implement incorporate emerge innovation promote support resource meet broad business goal	3.85

TASK 04

AUTOMATED TEXT SUMMARIZATION AND EVALUATION WITH SPACY AND ROUGE METRICS

spaCy: Text processing and tokenization.
TF-IDF: Quantify word importance in sentences.
Named Entity Recognition (NER): Enhance summarization with entity importance.

Process

- Lemmatize and tokenize document sentences.
- Calculate TF-IDF scores for words and named entities.
- Assign weights to sentences using TF-IDF and NER.
- Extract top-weighted sentences for the summary.

```
+------------------------------------------------------------------------------------------------------------------------------------------
| Summary
+------------------------------------------------------------------------------------------------------------------------------------------
| new Garmin 255w easy Set Accurate Directions location User Friendly Unit vehicle try function accurate leave battery mode stop Cracker Barrell lunch
| play trangle game tee Garmin load accurate map generally know road remotest area 0 5 star GPS Navigator navigate accurately straight road find Garmin
| software provide accurate direction whereever intend little disappointed inaccuracy post speed limit guilty pay close attention sign especially w
| interstate speed trap constantly change find map inaccurate update Garmin website golden accuracy yard day buy morning soon come ready navigate
| downfall product reason 5 star fact speed limit display road 100 accurate 255w good find street address business point interest hospital airport turn
| turn direction amazing accuracy
+------------------------------------------------------------------------------------------------------------------------------------------

gold_summary_1:
 --------------
Rouge-1: {'r': 0.13636363636363635, 'p': 0.03260869565217391, 'f': 0.052631575832564045}
Rouge-2: {'r': 0.0, 'p': 0.0, 'f': 0.0}
Rouge-l: {'r': 0.13636363636363635, 'p': 0.03260869565217391, 'f': 0.052631575832564045}

gold_summary_2:
 --------------
Rouge-1: {'r': 0.26666666666666666, 'p': 0.043478260869565216, 'f': 0.0747663527294961}
Rouge-2: {'r': 0.0, 'p': 0.0, 'f': 0.0}
Rouge-l: {'r': 0.26666666666666666, 'p': 0.043478260869565216, 'f': 0.0747663527294961}

gold_summary_3:
 --------------
Rouge-1: {'r': 0.0, 'p': 0.0, 'f': 0.0}
Rouge-2: {'r': 0.0, 'p': 0.0, 'f': 0.0}
Rouge-l: {'r': 0.0, 'p': 0.0, 'f': 0.0}

gold_summary_4:
 --------------
Rouge-1: {'r': 0.14285714285714285, 'p': 0.021739130434782608, 'f': 0.03773584676397309}
Rouge-2: {'r': 0.0, 'p': 0.0, 'f': 0.0}
Rouge-l: {'r': 0.14285714285714285, 'p': 0.021739130434782608, 'f': 0.03773584676397309}

gold_summary_5:
 --------------
Rouge-1: {'r': 0.07692307692307693, 'p': 0.010869565217391304, 'f': 0.019047616878004783}
Rouge-2: {'r': 0.0, 'p': 0.0, 'f': 0.0}
Rouge-l: {'r': 0.07692307692307693, 'p': 0.010869565217391304, 'f': 0.019047616878004783}
```

Evaluation

- Compare generated summary to reference summaries.
- Use ROUGE-1, ROUGE-2, and ROUGE-L scores for evaluation.

```
+-----------------------------------------------------------------------------------------------------------------------------------------------------------
| Summary
+-----------------------------------------------------------------------------------------------------------------------------------------------------------
| new Garmin 255w easy Set Accurate Directions location User Friendly Unit vehicle try function accurate leave battery mode stop Cracker Barrell lunch
| play trangle game tee Garmin load accurate map generally know road remotest area 0 5 star GPS Navigator navigate accurately straight road find Garmin
| software provide accurate direction whereever intend little disappointed inaccuracy post speed limit guilty pay close attention sign especially w
| interstate speed trap constantly change find map inaccurate update Garmin website golden accuracy yard day buy morning soon come ready navigate
| downfall product reason 5 star fact speed limit display road 100 accurate 255w good find street address business point interest hospital airport turn
| turn direction amazing accuracy
+-----------------------------------------------------------------------------------------------------------------------------------------------------------

gold_summary_1:
 ---------------
Rouge-1: {'r': 0.13636363636363635, 'p': 0.03260869565217391, 'f': 0.052631575832564045}
Rouge-2: {'r': 0.0, 'p': 0.0, 'f': 0.0}
Rouge-l: {'r': 0.13636363636363635, 'p': 0.03260869565217391, 'f': 0.052631575832564045}

gold_summary_2:
 ---------------
Rouge-1: {'r': 0.26666666666666666, 'p': 0.043478260869565216, 'f': 0.0747663527294961}
Rouge-2: {'r': 0.0, 'p': 0.0, 'f': 0.0}
Rouge-l: {'r': 0.26666666666666666, 'p': 0.043478260869565216, 'f': 0.0747663527294961}

gold_summary_3:
 ---------------
Rouge-1: {'r': 0.0, 'p': 0.0, 'f': 0.0}
Rouge-2: {'r': 0.0, 'p': 0.0, 'f': 0.0}
Rouge-l: {'r': 0.0, 'p': 0.0, 'f': 0.0}

gold_summary_4:
 ---------------
Rouge-1: {'r': 0.14285714285714285, 'p': 0.021739130434782608, 'f': 0.03773584676397309}
Rouge-2: {'r': 0.0, 'p': 0.0, 'f': 0.0}
Rouge-l: {'r': 0.14285714285714285, 'p': 0.021739130434782608, 'f': 0.03773584676397309}

gold_summary_5:
 ---------------
Rouge-1: {'r': 0.07692307692307693, 'p': 0.010869565217391304, 'f': 0.019047616878004783}
Rouge-2: {'r': 0.0, 'p': 0.0, 'f': 0.0}
Rouge-l: {'r': 0.07692307692307693, 'p': 0.010869565217391304, 'f': 0.019047616878004783}
```

TASK 05

ADVANCED TEXT SUMMARIZATION WITH TF-IDF, NAMED ENTITY RECOGNITION, AND SEMANTIC SIMILARITY

Objective
Improve summarization by minimizing redundancy among sentences.

Approach

- Use sentence-to-sentence semantic similarity
- Select top 20 sentences based on Sweight scores
- From top 20, select 10 sentences ensuring diversity
- Start with the highest Sweight sentence
- For the next, pick the one with the least similarity to the first, and so on
- Combine similarities for multi-sentence comparisons
- Stop once 10 diverse sentences are selected

Tools & Techniques

- **spaCy**: Semantic similarity, text processing, and NER
- **TF-IDF**: Quantify word and named entity importance

Summary
new Garmin 255w easy Set Accurate Directions location User Friendly Unit vehicle try accuracy yard lot friend address inaccurate GPS 0 5 star GPS
Navigator navigate accurately straight road glad buy like easy read graphic voice tell street turn uncannily accurate estimate mileage time arrival
destination update late 2010 map soon receive unit map accurate function accurate leave battery mode stop Cracker Barrell lunch play trangle game tee
direction provide accurate far accurate small glitch find explain con accurate determe original direction recalculate necessary highly accurate poi
great find map inaccurate update Garmin website golden

TASK 06

ADVANCED TEXT SUMMARIZATION WITH TF-IDF, NAMED ENTITIES, AND RAKE

Objective
Shift from histogram frequency to RAKE for keyword choice.

Tool
RAKE (Rapid Automatic Keyword Extraction)

Approach
- Preprocess and tokenize the document.
- Compute TF-IDF for sentences & named entities.
- Extract top 10 keywords using RAKE.
- Calculate sentence weights: TF-IDF, named entity importance, position-based scores, and RAKE keyword relevance.
- Select top 10 sentences based on weights.

Functionality
- Uses spaCy for text processing.
- RAKE for keyword extraction.
- TF-IDF for feature representation.

```
+-------------------------------------------------------------------------------------------------------------------------------------------------------------+
| Summary                                                                                                                                                     |
+-------------------------------------------------------------------------------------------------------------------------------------------------------------+
| new Garmin 255w easy Set Accurate Directions location User Friendly Unit vehicle try function accurate leave battery mode stop Cracker Barrell lunch      |
| play trangle game tee Garmin load accurate map generally know road remotest area 0 5 star GPS Navigator navigate accurately straight road find Garmin     |
| software provide accurate direction whereever intend little disappointed inaccuracy post speed limit guilty pay close attention sign especially w          |
| interstate speed trap constantly change find map inaccurate update Garmin website golden accuracy yard day buy morning soon come ready navigate           |
| downfall product reason 5 star fact speed limit display road 100 accurate provide immediate alternative route online map program inaccurate block          |
| obstacle                                                                                                                                                    |
+-------------------------------------------------------------------------------------------------------------------------------------------------------------+
```

TASK 07

DOCUMENT SUMMARIZATION USING VARIOUS SUMY ALGORITHMS

Objective
Test multiple summarization algorithms on the same document.

Tool
Sumy

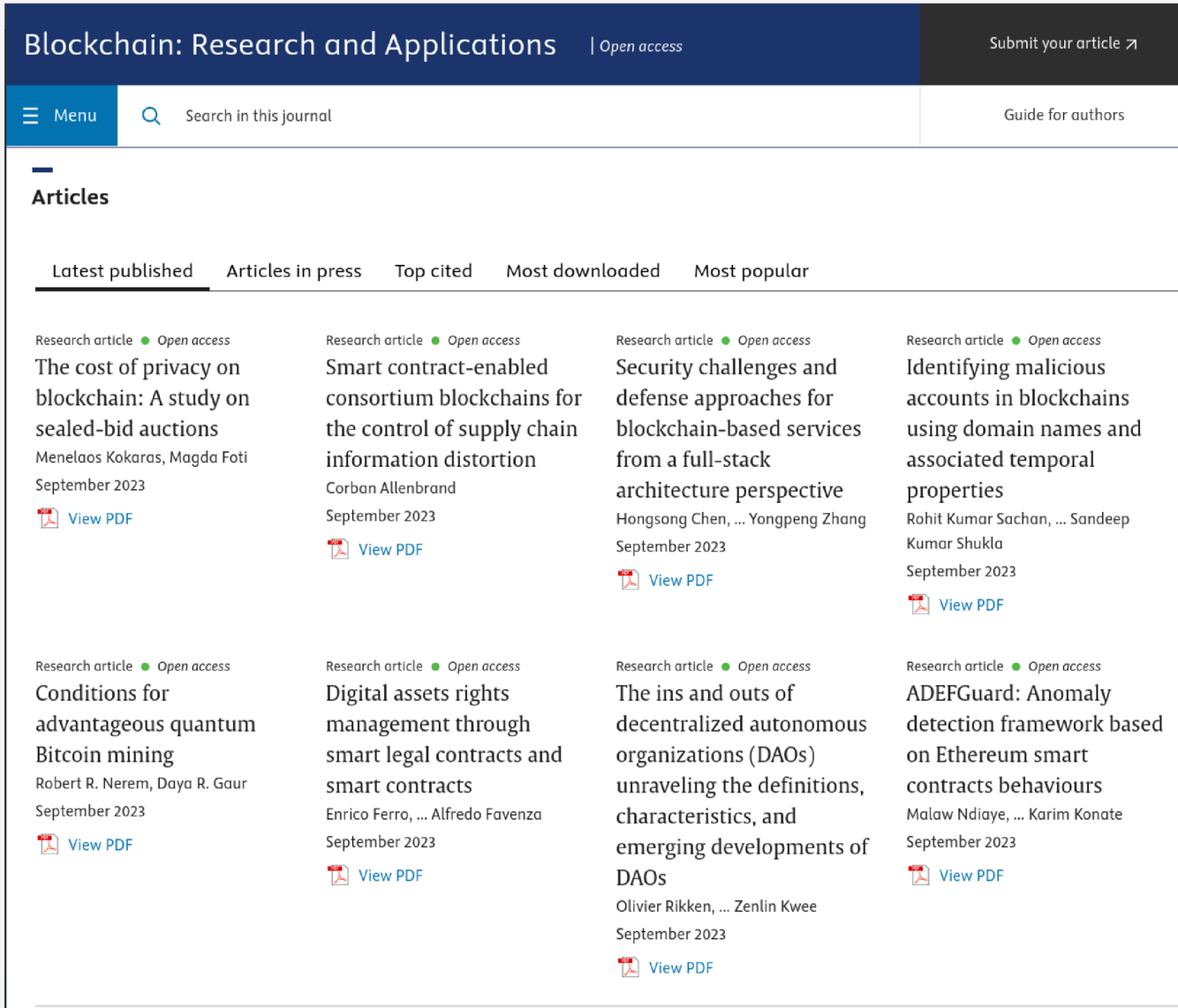
Method
- Load the desired document.
- Choose a Sumy summarizer.
- Generate the summary with a defined number of sentences (SENTENCES_COUNT = 10).
- Display summary using PrettyTable.

Summaries generated by the following algorithms:
- LsaSummarizer
- LuhnSummarizer
- EdmundsonSummarizer
- LexRankSummarizer
- TextRankSummarizer
- SumBasicSummarizer
- KLSummarizer

DOCUMENT SUMMARIZATION USING VARIOUS SUMY ALGORITHMS

TASK 08

COMPARISON OF SUMMARIZATION TECHNIQUES ON A NEW DATASET



Blockchain: Research and Applications | Open access

Submit your article

Menu | Search in this journal | Guide for authors

Articles

Latest published | Articles in press | Top cited | Most downloaded | Most popular

Research article • Open access
The cost of privacy on blockchain: A study on sealed-bid auctions
Menelaos Kokaras, Magda Foti
September 2023
View PDF

Research article • Open access
Smart contract-enabled consortium blockchains for the control of supply chain information distortion
Corban Allenbrand
September 2023
View PDF

Research article • Open access
Security challenges and defense approaches for blockchain-based services from a full-stack architecture perspective
Hongsong Chen, ... Yongpeng Zhang
September 2023
View PDF

Research article • Open access
Identifying malicious accounts in blockchains using domain names and associated temporal properties
Rohit Kumar Sachan, ... Sandeep Kumar Shukla
September 2023
View PDF

Research article • Open access
Conditions for advantageous quantum Bitcoin mining
Robert R. Nerem, Daya R. Gaur
September 2023
View PDF

Research article • Open access
Digital assets rights management through smart legal contracts and smart contracts
Enrico Ferro, ... Alfredo Favenza
September 2023
View PDF

Research article • Open access
The ins and outs of decentralized autonomous organizations (DAOs) unraveling the definitions, characteristics, and emerging developments of DAOs
Olivier Rikken, ... Zenlin Kwee
September 2023
View PDF

Research article • Open access
ADEFGuard: Anomaly detection framework based on Ethereum smart contracts behaviours
Malaw Ndiaye, ... Karim Konate
September 2023
View PDF

Objective
Evaluate the performance of various summarization techniques on a set of scientific papers from an Elsevier journal.

Dataset
- Choose an Elsevier journal.
- Select 10 high-citation papers.
- Use the Introduction as the main document for summarization.
- Use Abstract and Conclusion as reference summaries for evaluation.

Evaluation Metrics
1. ROUGE-1
2. ROUGE-2

Summarization Techniques
- TF-IDF and POS Weights in spaCy (Task 03)
- Advanced Text Summarization with TF-IDF, NER, and Semantic Similarity (Task 05)
- Various Sumy Algorithms (Task 07):
  - LSA Summarizer
  - Luhn Summarizer
  - Edmundson Summarizer
  - LexRank Summarizer
  - TextRank Summarizer
  - SumBasic Summarizer
  - KL-Sum Summarizer

Implementation:
- For each paper, generate a summary using each technique.
- Compute the ROUGE scores comparing the generated summary against the reference summaries (Abstract and Conclusion).
- Present the results in a table format for comparison.

ROUGE-1 AND ROUGE-2 EVALUATION

JOURNAL - Blockchain: Research and Applications https://www.sciencedirect.com/journal/blockchain-research-and-applications

Top cited papers

1 - Blockchain technology applications for Industry 4.0: A literature-based review
https://www.sciencedirect.com/science/article/pii/S2096720921000221

Golden Summary : /content/task-08/1-s2.0-main.gold

Main Document : /content/task-08/1-s2.0-main.txt.data

2 - A survey on the adoption of blockchain in IoT: challenges and solutions
https://www.sciencedirect.com/science/article/pii/S2096720921000014

Golden Summary : /content/task-08/2-s2.0-main.gold

Main Document : /content/task-08/2-s2.0-main.txt.data

3 - A survey on blockchain technology and its security https://www.sciencedirect.com/science/article/pii/S2096720922000070

Golden Summary : /content/task-08/3-s2.0-main.gold

Main Document : /content/task-08/3-s2.0-main.txt.data

4 - ABCDE—agile block chain DApp engineering https://www.sciencedirect.com/science/article/pii/S2096720920300026

Golden Summary : /content/task-08/4-s2.0-main.gold

Main Document : /content/task-08/4-s2.0-main.txt.data

5 - The case of HyperLedger Fabric as a blockchain solution for healthcare applications
https://www.sciencedirect.com/science/article/pii/S2096720921000075

Golden Summary : /content/task-08/5-s2.0-main.gold

Main Document : /content/task-08/5-s2.0-main.txt.data

6 - Applications of Blockchain Technology in marketing—A systematic review of marketing technology companies
https://www.sciencedirect.com/science/article/pii/S209672092100018X

Task 03 - Document Summarization Using TF-IDF and POS Weights in spaCy:

generated_summary_ts3_ts8_summary
article address follow research question RQ1 study major benefit blockchain context industry 4.0 RQ2 identify stud
4.0 RQ3 study associate blockchain capability successful industry 4.0 implementation perspective RQ4 study differe
realisation rq5 identify study major application Blockchain technology industry 4.0 datum analytic Artificial Inte
information previously store Enterprise Resource Planning ERP company paper aim study significant potential role B
identify potential role Blockchain Industry 4.0 supply chain dynamic structure comprise multiple business work sat
product Blockchain Blockchain define decentralised distribute directory drive smart contract provide opportunity t
chain payment application business transaction main chain support secondary tertiary chain co operate form ecosyst
distribution strategy implement incorporate emerge innovation promote support resource meet broad business goal

Metric	Recall	Precision	F-Score
rouge-1	0.1057	0.2393	0.1466
rouge-2	0.0140	0.0414	0.0209
rouge-l	0.1019	0.2308	0.1414

Task 05 - Advanced Text Summarization with TF-IDF, Named Entity Recognition, and Semantic Similarity:

generated_summary_ts5_ts8_
article address follow research question RQ1 study major benefit blockchain context industry 4.0 RQ2 identify stud
4.0 RQ3 study associate blockchain capability successful industry 4.0 implementation perspective RQ4 study differe
realisation rq5 identify study major application Blockchain technology industry 4.0 instance Leng et al Blockchain
computer allow tamper proof real time log create main chain support secondary tertiary chain co operate form ecosy
Artificial Intelligence AI understand case application foreign currency fiat currency problem exclude control supp
potential role Blockchain Industry 4.0 database run network computer call node single point failure information ac
overcome potential cybersecurity barrier achieve intelligence Industry 4.0 process take information previously sto

Metric	Recall	Precision	F-Score
rouge-1	0.1057	0.2500	0.1485
rouge-2	0.0140	0.0438	0.0212
rouge-l	0.1019	0.2411	0.1432

generated_LuhnSummarizer
In the current scenario, it is necessary to understand blockchain and its value for the effective implementation of In
comprising multiple businesses that work together to satisfy customers' needs by adding value from the raw material le
following research questions: RQ1: To study major benefits of blockchain in the context of Industry 4.0; RQ2: To ident
technology for industry 4.0; RQ3: To study associated blockchain capabilities for successful Industry 4.0 implementati
4.0 spheres/sub-domains for Blockchain technology realisation; RQ5: To identify and study major applications of Blockc
In Industry 4.0 environment, real-time information is needed to create a smooth manufacturing and service system. So,
role of Blockchain in Industry 4.0. In this study we have focused on various drivers, enablers, and associated capabil
are holistic in nature. Blockchain Blockchain can be defined as a decentralised, distributed directory driving smart c
traceability aid, record management, automation for the supply chain, payment applications and other business transact
blockchain, other than the nodes that make it superior to other data storage technologies. Blockchain technology, some
relatively new form of a database for transaction information stored in a decentralised and transparent manner. The da
nodes, so there is no single-point-of-failure, and information can be accessed in real-time.

Metric	Recall	Precision	F-Score
rouge-1	0.2189	0.3671	0.2742
rouge-2	0.0818	0.1367	0.1023
rouge-l	0.2075	0.3481	0.2600

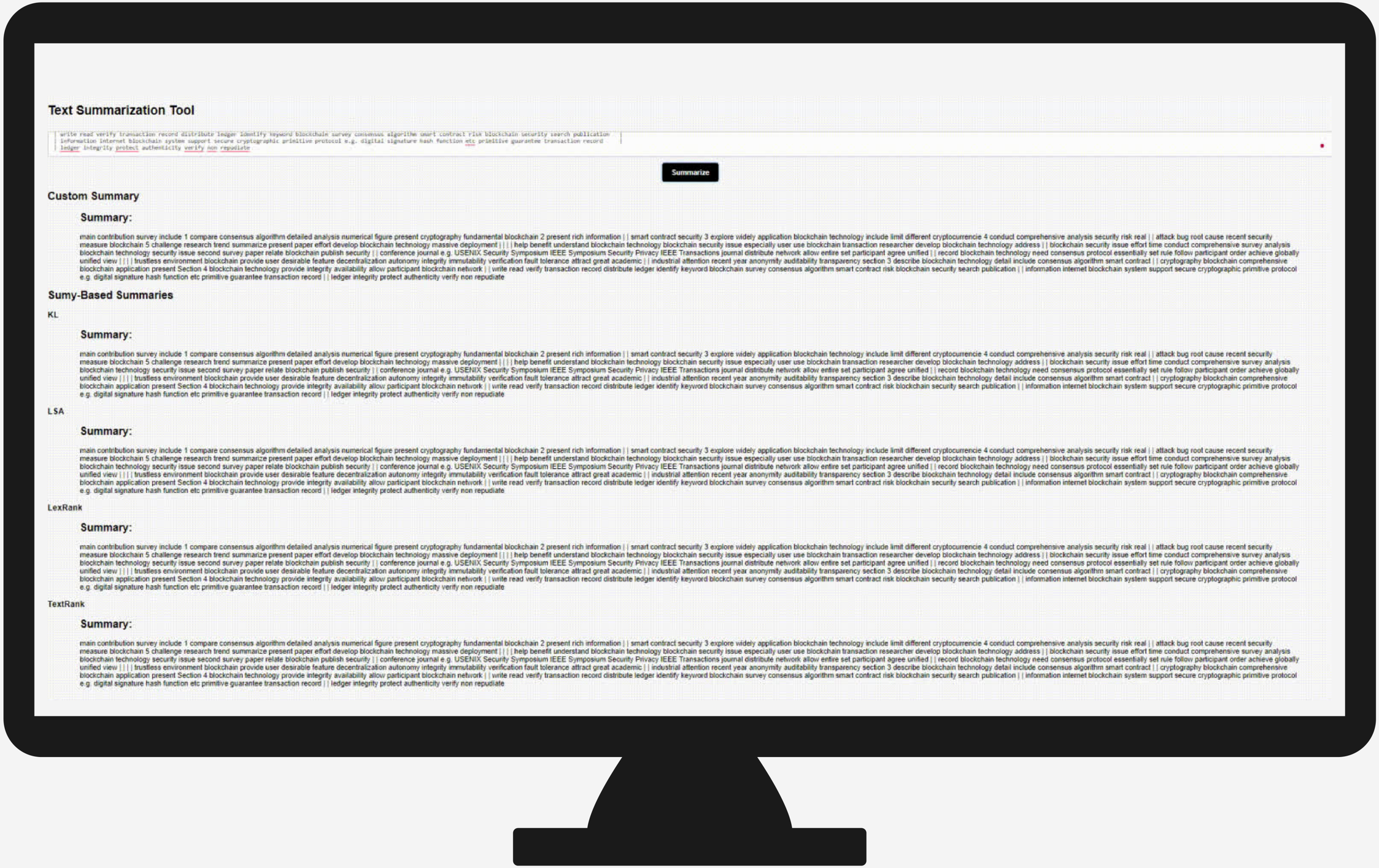
Edmundson Summarizer:

generated_EdmundsonSummarizer
In the current scenario, it is necessary to understand blockchain and its value for the effective implementation of In
for blockchain, like financial transactions applications in which blockchains can provide trust. Foreign currencies ar
controlled supply transaction may take place. The product itself and its assembly's identification part can also be li
It provides a reminder where the ability to recognise goods with the defect may be beneficial. Here, blockchain will p
assemblies, parts, sales paths, etc. This paper aims to study the significant potential role of Blockchain for Industr
questions: RQ1: To study major benefits of blockchain in the context of Industry 4.0; RQ2: To identify and study major
industry 4.0; RQ3: To study associated blockchain capabilities for successful Industry 4.0 implementation perspectives
domains for Blockchain technology realisation; RQ5: To identify and study major applications of Blockchain technology
information, blockchain is the perfect technology which can fulfil major challenges. It allows users to preserve their
their permission that nobody can read or view.

Metric	Recall	Precision	F-Score
rouge-1	0.2151	0.4071	0.2815
rouge-2	0.0537	0.1127	0.0728
rouge-l	0.2000	0.3786	0.2617

TASK 09
DESIGN SIMPLE GUI

spaCy


Text Summarization Tool
write read verify transaction record distribute ledger identify keyword blockchain survey consensus algorithm smart contract risk blockchain security search publication information internet blockchain system support secure cryptographic primitive protocol e.g. digital signature hash function etc primitive guarantee transaction record ledger integrity protect authenticity verify non repudiate
Summarize
Custom Summary
Summary:
main contribution survey include 1 compare consensus algorithm detailed analysis numerical figure present cryptography fundamental blockchain 2 present rich information | | smart contract security 3 explore widely application blockchain technology include limit different cryptocurrencie 4 conduct comprehensive analysis security risk real | | attack bug root cause recent security measure blockchain 5 challenge research trend summarize present paper effort develop blockchain technology massive deployment | | | | help benefit understand blockchain technology blockchain security issue especially user use blockchain transaction researcher develop blockchain technology address | | blockchain security issue effort time conduct comprehensive survey analysis blockchain technology security issue second survey paper relate blockchain publish security | | conference journal e.g. USENIX Security Symposium IEEE Symposium Security Privacy IEEE Transactions journal distribute network allow entire set participant agree unified | | record blockchain technology need consensus protocol essentially set rule follow participant order achieve globally unified view | | | | trustless environment blockchain provide user desirable feature decentralization autonomy integrity immutability verification fault tolerance attract great academic | | industrial attention recent year anonymity auditability transparency section 3 describe blockchain technology detail include consensus algorithm smart contract | | cryptography blockchain comprehensive blockchain application present Section 4 blockchain technology provide integrity availability allow participant blockchain network | | write read verify transaction record distribute ledger identify keyword blockchain survey consensus algorithm smart contract risk blockchain security search publication | | information internet blockchain system support secure cryptographic primitive protocol e.g. digital signature hash function etc primitive guarantee transaction record | | ledger integrity protect authenticity verify non repudiate
Sumy-Based Summaries
KL
Summary:
main contribution survey include 1 compare consensus algorithm detailed analysis numerical figure present cryptography fundamental blockchain 2 present rich information | | smart contract security 3 explore widely application blockchain technology include limit different cryptocurrencie 4 conduct comprehensive analysis security risk real | | attack bug root cause recent security measure blockchain 5 challenge research trend summarize present paper effort develop blockchain technology massive deployment | | | | help benefit understand blockchain technology blockchain security issue especially user use blockchain transaction researcher develop blockchain technology address | | blockchain security issue effort time conduct comprehensive survey analysis blockchain technology security issue second survey paper relate blockchain publish security | | conference journal e.g. USENIX Security Symposium IEEE Symposium Security Privacy IEEE Transactions journal distribute network allow entire set participant agree unified | | record blockchain technology need consensus protocol essentially set rule follow participant order achieve globally unified view | | | | trustless environment blockchain provide user desirable feature decentralization autonomy integrity immutability verification fault tolerance attract great academic | | industrial attention recent year anonymity auditability transparency section 3 describe blockchain technology detail include consensus algorithm smart contract | | cryptography blockchain comprehensive blockchain application present Section 4 blockchain technology provide integrity availability allow participant blockchain network | | write read verify transaction record distribute ledger identify keyword blockchain survey consensus algorithm smart contract risk blockchain security search publication | | information internet blockchain system support secure cryptographic primitive protocol e.g. digital signature hash function etc primitive guarantee transaction record | | ledger integrity protect authenticity verify non repudiate
LSA
Summary:
main contribution survey include 1 compare consensus algorithm detailed analysis numerical figure present cryptography fundamental blockchain 2 present rich information | | smart contract security 3 explore widely application blockchain technology include limit different cryptocurrencie 4 conduct comprehensive analysis security risk real | | attack bug root cause recent security measure blockchain 5 challenge research trend summarize present paper effort develop blockchain technology massive deployment | | | | help benefit understand blockchain technology blockchain security issue especially user use blockchain transaction researcher develop blockchain technology address | | blockchain security issue effort time conduct comprehensive survey analysis blockchain technology security issue second survey paper relate blockchain publish security | | conference journal e.g. USENIX Security Symposium IEEE Symposium Security Privacy IEEE Transactions journal distribute network allow entire set participant agree unified | | record blockchain technology need consensus protocol essentially set rule follow participant order achieve globally unified view | | | | trustless environment blockchain provide user desirable feature decentralization autonomy integrity immutability verification fault tolerance attract great academic | | industrial attention recent year anonymity auditability transparency section 3 describe blockchain technology detail include consensus algorithm smart contract | | cryptography blockchain comprehensive blockchain application present Section 4 blockchain technology provide integrity availability allow participant blockchain network | | write read verify transaction record distribute ledger identify keyword blockchain survey consensus algorithm smart contract risk blockchain security search publication | | information internet blockchain system support secure cryptographic primitive protocol e.g. digital signature hash function etc primitive guarantee transaction record | | ledger integrity protect authenticity verify non repudiate
LexRank
Summary:
main contribution survey include 1 compare consensus algorithm detailed analysis numerical figure present cryptography fundamental blockchain 2 present rich information | | smart contract security 3 explore widely application blockchain technology include limit different cryptocurrencie 4 conduct comprehensive analysis security risk real | | attack bug root cause recent security measure blockchain 5 challenge research trend summarize present paper effort develop blockchain technology massive deployment | | | | help benefit understand blockchain technology blockchain security issue especially user use blockchain transaction researcher develop blockchain technology address | | blockchain security issue effort time conduct comprehensive survey analysis blockchain technology security issue second survey paper relate blockchain publish security | | conference journal e.g. USENIX Security Symposium IEEE Symposium Security Privacy IEEE Transactions journal distribute network allow entire set participant agree unified | | record blockchain technology need consensus protocol essentially set rule follow participant order achieve globally unified view | | | | trustless environment blockchain provide user desirable feature decentralization autonomy integrity immutability verification fault tolerance attract great academic | | industrial attention recent year anonymity auditability transparency section 3 describe blockchain technology detail include consensus algorithm smart contract | | cryptography blockchain comprehensive blockchain application present Section 4 blockchain technology provide integrity availability allow participant blockchain network | | write read verify transaction record distribute ledger identify keyword blockchain survey consensus algorithm smart contract risk blockchain security search publication | | information internet blockchain system support secure cryptographic primitive protocol e.g. digital signature hash function etc primitive guarantee transaction record | | ledger integrity protect authenticity verify non repudiate
TextRank
Summary:
main contribution survey include 1 compare consensus algorithm detailed analysis numerical figure present cryptography fundamental blockchain 2 present rich information | | smart contract security 3 explore widely application blockchain technology include limit different cryptocurrencie 4 conduct comprehensive analysis security risk real | | attack bug root cause recent security measure blockchain 5 challenge research trend summarize present paper effort develop blockchain technology massive deployment | | | | help benefit understand blockchain technology blockchain security issue especially user use blockchain transaction researcher develop blockchain technology address | | blockchain security issue effort time conduct comprehensive survey analysis blockchain technology security issue second survey paper relate blockchain publish security | | conference journal e.g. USENIX Security Symposium IEEE Symposium Security Privacy IEEE Transactions journal distribute network allow entire set participant agree unified | | record blockchain technology need consensus protocol essentially set rule follow participant order achieve globally unified view | | | | trustless environment blockchain provide user desirable feature decentralization autonomy integrity immutability verification fault tolerance attract great academic | | industrial attention recent year anonymity auditability transparency section 3 describe blockchain technology detail include consensus algorithm smart contract | | cryptography blockchain comprehensive blockchain application present Section 4 blockchain technology provide integrity availability allow participant blockchain network | | write read verify transaction record distribute ledger identify keyword blockchain survey consensus algorithm smart contract risk blockchain security search publication | | information internet blockchain system support secure cryptographic primitive protocol e.g. digital signature hash function etc primitive guarantee transaction record | | ledger integrity protect authenticity verify non repudiate

LINKS

GOOGLE COLAB LINK

GITHUB LINK

THANK YOU!