MENEMUKAN KEBAHARUAN PENELITIAN (Novelty)

Dr. Jasmir, S.Kom., M.Kom

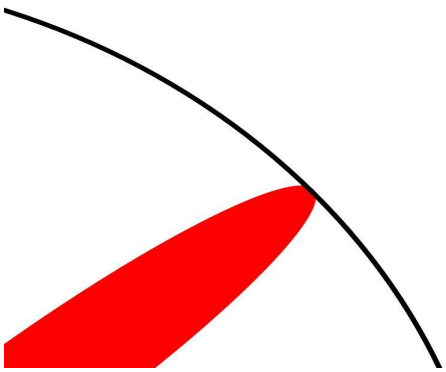
26 Januari 2023
Universitas Dinamika Bangsa

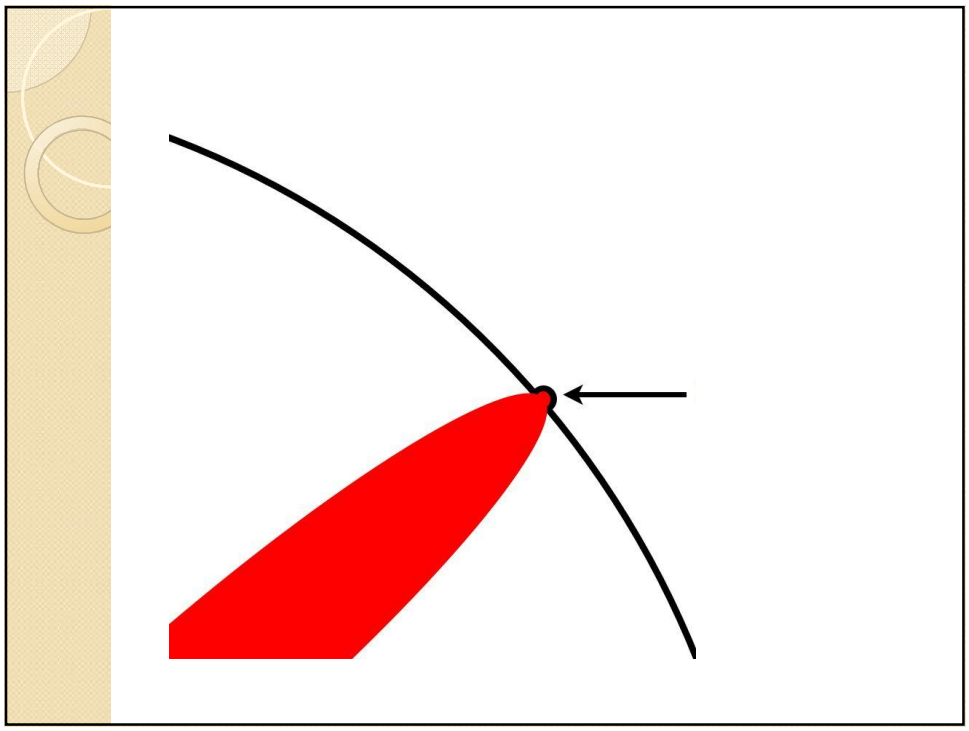
Apa itu Novelty

Pertanyaannya simple, jawabannya tidak simple

- Novelty adalah **unsur kebaruan atau temuan dari sebuah penelitian**. Penelitian dikatakan baik jika menemukan unsur temuan baru sehingga memiliki kontribusi baik bagi keilmuan maupun bagi kehidupan.
- Novelty atau kebaruan atau nilai tambah dalam riset ilmiah adalah mutlak.
- Tanpa novelty, anda layaknya mengulang apa yang sebenarnya telah dilakukan orang lain.

- Ada 2 Kontribusi (bidang IT)
  - Kontribusi pada **Metode**
  - Kontribusi pada **Permasalahan**





Secara khusus

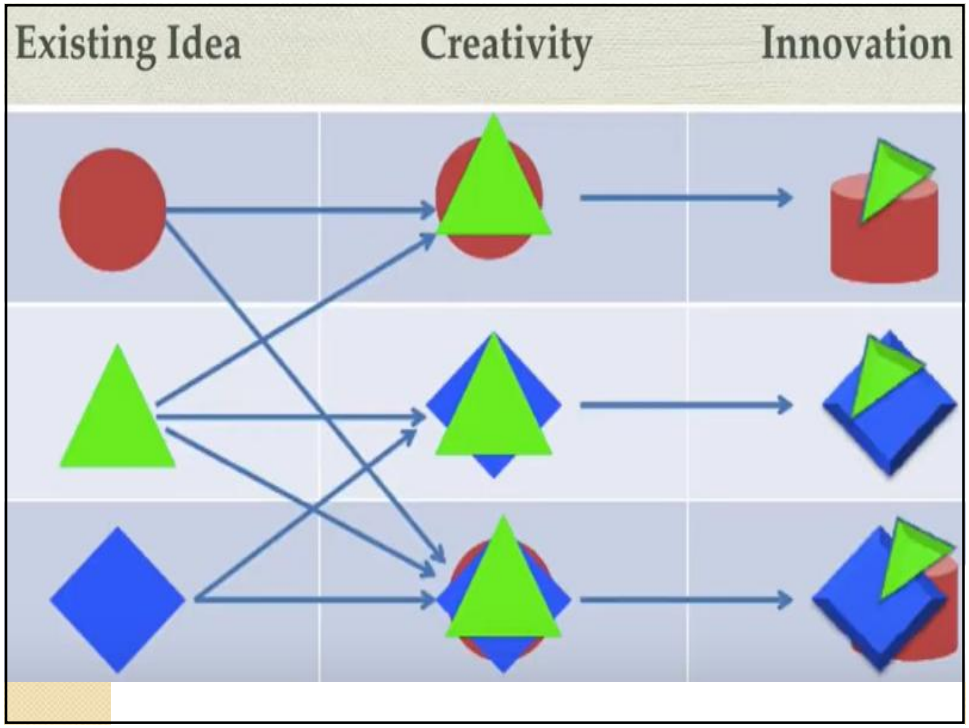
- Ciri Khas kebaharuan adalah dalam bidang IT yaitu ***"Improvement"***
  - Optimasi (min atau maks)
  - Optimalisasi (paling baik atau paling tinggi)
  - Otomatisasi
  - Reduksi
  - Ekstraksi
  - Efisiensi
  - Dan lain lain

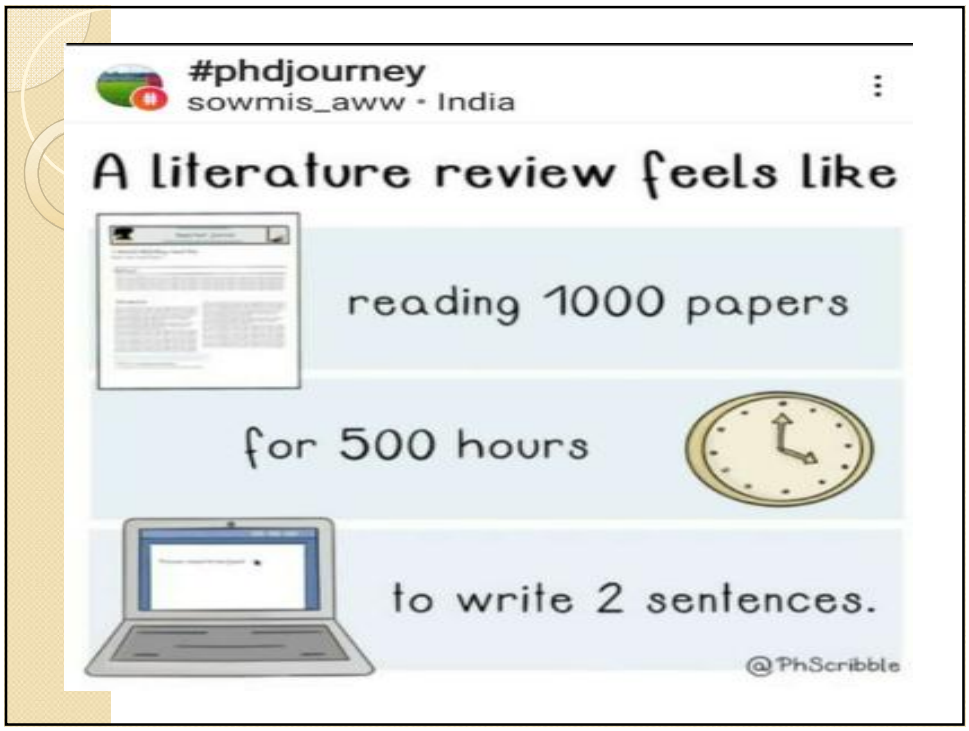
Secara Umum (bidang IT)

- Working on old problem with new method (improvement-type research, **contribution in method**)
- Implementing old method to new problem **(contribution in problem)**
- Implementing new method to new problem (combine 1 and 2), **(contribution in method and problem)**

- Proofing a theorical model (computational) with real implementation in hardware
- Development of low cost method/technique/hardware/system with equal performance of existing method/technique/hardware/system

- (yang termasuk) New Methods
  - Metode/Algoritma baru
  - Kombinasi 2 metode yang sudah ada
  - Kombinasi metode yang sudah ada dengan fitur yang sudah ada
- (yang termasuk) New Problem
  - Perbedaan pada disiplin ilmu
  - Perbedaan pada tema (dalam satu keilmuan)
  - Perbedaan pada fitur (dalam satu tema)





ABSTRACT Feature selection (FS) is one of the important tasks of data preprocessing in data analytics. The data with a large number of features will affect the computational complexity, increase a huge amount of resource usage and time consumption for data analytics. The objective of this study is to analyze relevant and significant features of huge network traffic to be used to improve the accuracy of traffic anomaly detection and to decrease its execution time. Information Gain is the most feature selection technique used in Intrusion Detection System (IDS) research. This study uses Information Gain, ranking and grouping the features according to the minimum weight values to select relevant and significant features, and then implements Random Forest (RF), Bayes Net (BN), Random Tree (RT), Naive Bayes (NB) and J48 classifier algorithms in experiments on CICIDS-2017 dataset. The experiment results show that the number of relevant and significant features yielded by Information Gain affects significantly the improvement of detection accuracy and execution time. Specifically, the Random Forest algorithm has the highest accuracy of 99.86% using the relevant selected features of 22, whereas the J48 classifier algorithm provides an accuracy of 99.87% using 52 relevant selected features with longer execution time.

CICIDS-2017 Dataset Feature Analysis With Information Gain for Anomaly Detection

KURNIABUDI[1,2], (Member, IEEE), DERIS STIAWAN[3], DARMAWIJOYO[4], MOHD YAZID BIN IDRIS[5], (Member, IEEE), ALWI M. BAMHDI[6], AND RAHMAT BUDIARTO[7]

ABSTRACT

The difficulty of the intrusion detection system in heterogeneous networks is significantly affected by devices, protocols, and services, thus the network becomes complex and difficult to identify. Deep learning is one algorithm that can classify data with high accuracy. In this research, we proposed deep learning to intrusion detection system identification methods in heterogeneous networks to increase detection accuracy. In this paper, we provide an overview of the proposed algorithm, with an initial experiment of denial of services (DoS) attacks and results. The results of the evaluation showed that deep learning can improve detection accuracy in the heterogeneous internet of things (IoT).

Intrusion detection with deep learning on internet of things heterogeneous network

Sharipuddin[1], Benni Purnama[2], Kurniabudi[3], Eko Arip Winanto[4], Deris Stiawan[5], Darmawijoyo Hanapi[6], Mohd. Yazid Idris[7], Rahmat Budiarto[8]

[1,2,3,4,5,6]Department of Computer Science, Universitas Sriwijaya, Indonesia
[1,2,3]Computer Enginering, Universitas Dinamika Bangsa, Jambi Indonesia
[4,7]School of Computing, Universiti Teknologi Malaysia, Malaysia
[8]College of Computer Science & IT, Albaha University, Saudi Arabia

ABSTRACT

Ransomware is able to attack and take over access of the targeted user's computer. Then the hackers demand a ransom to restore the user's access rights. Ransomware detection process especially in big data has problems in term of computational processing time or detection speed. Thus, it requires a dimensionality reduction method for computational process efficiency. This research work investigates the efficiency of three dimensionality reduction methods, i.e.: Principal Component Analysis (PCA), Factor Analysis (FA) and Truncated Singular Value Decomposition (TSVD). Experimental results on CICAndMal2017 dataset show that PCA is the fastest and most significant method in the computational process with average detection time of 34.33s. Furthermore, result of accuracy, precision and recall also show that the PCA is superior compared to FA and TSVD.

Time Efficiency on Computational Performance of PCA, FA and TSVD on Ransomware Detection

Benni Purnama[1], Deris Stiawan[2], Darmawijoyo Hanapi[3], Mohd Yazid Bin Idris[4], Sharipuddin[5], Nurul Afifah[6], Rahmat Budiarto[7]

[1,5]Department of Computer Engineering, Universitas Dinamika Bangsa Jambi, Indonesia
[2,6]Department of Computer Engineering, Universitas Sriwijaya, Indonesia [3]Faculty of Mathematics and Natural Science, Universitas Sriwijaya, Indonesia [4]School of Computing, Universiti Teknologi Malaysia, Malaysia
[7]College of Computer Science and IT, Albaha University, Albaha, Saudi Arabia

The positioning system research strongly supports the development of location-based services used by related business organizations. However, location-based services with user experience still have many obstacles to overcome, including how to maintain a high level of position accuracy. From the literature studies reviewed, it is necessary to develop an indoor positioning system using fingerprinting based on Received Signal Strength (RSS). So far, the testing of the indoor positioning system has been carried out with an algorithm. But, in this research, with the proposed parameters, we will conduct experiments with a learning approach. The data tested is the signal service data on the device in the Dinamika Bangsa University building. The test was conducted with a deep learning approach using a deep neural network (DNN) algorithm. The DNN method can estimate the actual space and get better position results, whereas machine learning methods such as the DNN algorithm can handle more effectively large data and produce more accurate data. From the results of comparative testing with the learning approach between DNN, KNN, and SVM, it can be concluded that the evaluation with KNN is slightly better than the use of DNN in a single case. However, the results of KNN have low consistency; this is seen from the fluctuations in the movements of the *R2* score and MSE values produced. Meanwhile, DNN gives a consistent value even though it has varied hidden layers. The Support Vector Machine (SVM) gives the worst value of these experiments, although, in the past, SVM was known as one of the favorite methods.

Indoor Positioning System in Learning Approach Experiments

Dodo Zaenal Abidin[1], **Siti Nurmaini**[2], **Erwin**[2], **Errissya Rasywir**[3], and **Yovi Pratama**[3]

[1]*Department of Engineering, Sriwijaya University, Palembang, Indonesia*
[2]*Department of Informatics Engineering, Sriwijaya University, Palembang, Indonesia*
[3]*Department of Informatics Engineering, Dinamika Bangsa University, Jambi, Indonesia*

Image retrieval using features has been used in previous studies including shape, color, texture, but these features are lagging. With the selection of high-level features with contours, this research is done with the hypothesis that images on objects can also be subjected to representations that are commonly used in natural images. Considering the above matters, we need to research the feature extraction of object images using gradation contour. From the results of the gradation contour test results, there is linearity between the results of accuracy with the large number of images tested. Therefore, it can be said that the influence of the number of images will affect the accuracy of classification. The use of contour gradation can be accepted and treated equally in all image types, so there is no more differentiation between image features. The complexity of the image does not affect the method of extracting features that are only used uniquely by an image. From the results of testing the polynomial coefficient savings data as a result of the gradation contour, the highest result is 81.40% with the highest number of categories and the number of images tested in the category is also higher.

Extraction of object image features with gradation contour

Fachruddin[1], Saparudin[2], Errissya Rasywir[3], Yovi Pratama[4], Beni Irawan[5]
[1]Department of Informatics Engineering, Faculty of Engineering, Universitas Sriwijaya, Palembang, Indonesia
[2]Department Informatics, Telkom University, Bandung, Indonesia
[3,4,5]Department of Informatics Engineering, Faculty of Computer Science, Universitas Dinamika Bangsa, Jambi, Indonesia

ABSTRACT

Augmented reality (AR) has been applied in maintenance, simulation, remote assistance, and other fields. One of the issues arising from these applications is how objects can be placed in physical environments using an AR system. Object placement consists of two processes: object detection and segmentation. Due to the importance of placement, in this paper, we propose using deep learning to address issues with the placement of objects through detection and segmentation in AR. Deep learning can help complete tasks by providing correct information about environmental changes in real-time situations. The problem is that it is rarely used in AR, which suggests a combination of deep learning-based object detection and instance segmentation with wearable AR technology to improve performance on complex tasks. In our work, we propose to address this problem by applying a convolutional neural network to facilitate the detection and segmentation of objects in real environments. To measure AR performance, we examined detection accuracy in environments with different intensities. The results of the experiment demonstrate satisfactory performance, reaching 98% for segmentation and accurate detection.

Enhancing Object Tracking in Augmented Reality Using Convolutional Neural Network-Based

Nurhadi[1], Deris Stiawan[2], Mohd. Yazid Idris[3], Saparudin[4]
[1]Department of Informatics Engineering, Dinamika Bangsa University; Faculty of Engineering, Sriwijaya University, Indonesia
[2]Department of Computer Engineering, Universitas Sriwijaya, Indonesia
[3]School of Computing, Faculty of Engineering, Universiti Teknologi Malaysia
[4]Department of Informatics Engineering, Telkom University, Indonesia

Abstract: Text classification is an important component in many applications. Text classification has attracted the attention of researchers to continue to develop innovations and build new classification models that are sourced from clinical trial texts. In building classification models, many methods are used, including supervised learning. The purpose of this study is to improve the computational performance of one of the supervised learning methods, namely KNN, in building a clinical trial document text classification model by combining KNN and the fine-grained algorithm. This research contributed to increasing the computational performance of KNN from 388,274 s to 260,641 s in clinical trial texts on a clinical trial text dataset with a total of 1,000,000 data.

Fine-Grained Algorithm for Improving KNN Computational Performance on Clinical Trials Text Classification

Jasmir Jasmir [1,2,*], Siti Nurmaini [3] and Bambang Tutuko [3]

[1] Doctoral Program of Engineering Science, Faculty of Engineering, Universitas Sriwijaya, Palembang 30128, Indonesia
[2] Computer Engineering, Universitas Dinamika Bangsa, Jambi 36138, Indonesia
[3] Intelligent System Research Group, Universitas Sriwijaya, Palembang 30128, Indonesia; sitinurmaini@gmail.com (S.N.); bambang_tutuko@unsri.ac.id (B.T.)
* Correspondence: ijay_jasmir@yahoo.com

applied sciences

MDPI

Article

Learning Eligibility in Cancer Clinical Trials Using Deep Neural Networks

Aurelia Bustos and Antonio Pertusa *

Pattern Recognition and Artificial Intelligence Group (GRFIA), Department of Software and Computing Systems, University Institute for Computing Research, University of Alicante, E-03690 Alicante, Spain; aurelia@medbravo.org
* Correspondence: pertusa@dlsi.ua.es

Table 3. Overall results on the validation set for all the classifiers using a dataset of 10^6 samples and the full dataset (4.1×10^6) samples. Both experiments were performed using 20% of the prototypes for validation and 80% for training. The best results are marked in bold.

Classifier	Dataset Size	Precision	Recall	F_1	Cohen's κ
FastText	10^6	0.88	0.86	0.87	0.75
	4.1×10^6	0.89	0.87	0.88	0.76
CNN	10^6	0.88	0.88	0.88	0.76
	4.1×10^6	0.91	0.91	0.91	0.83
SVM	10^6	0.79	0.79	0.79	0.57
	4.1×10^6	0.79	0.79	0.79	0.58
kNN	10^6	0.92	0.92	0.92	0.83
	4.1×10^6	0.93	0.93	**0.93**	**0.84**

5. Conclusions

In this work, we have trained, validated, and compared various classifiers (FastText and a CNN with pre-trained word-embeddings, kNN, and SVM) on a corpus of cancer clinical trial protocols (www.clinicaltrials.gov). The models classify short free-text sentences describing clinical information (medical history, concomitant medication, type and features of tumor, such as molecular profile, cancer therapy, etc.) as eligible or not eligible criteria for volunteering in these trials. SVM yielded the lowest accuracy results, and kNN obtained top accuracy performance similar to the CNN model, but it had the lowest computational performance. Particularly, the high accuracy achieved with kNN is the immediate consequence of using as input a highly efficient clinical statement representation which is based on averaged pre-trained word-embeddings. A possible reason for this is that the kNN accuracy relies almost exclusively on using a highly efficient vector representation as the input data and on the dataset size. Being a non-parametric method, it is often successful—as in this case—in classification situations where the decision boundary is very irregular. Nonetheless, in spite of its high accuracy and the minimal training phase, we favor the use of deep learning architectures for classification (such as CNN) over a kNN model because of its lower computational cost during prediction time. In fact, classifying a given observation requires a rundown of the whole dataset being too computationally expensive for large dataset as in this work.

Social Network Analysis and Mining (2018) 8:29
https://doi.org/10.1007/s13278-018-0508-z

ORIGINAL ARTICLE

Fine-grained document clustering via ranking and its application to social media analytics

Taufik Sutanto[1] · Richi Nayak[2]

Abstract

Extracting valuable insights from a large volume of unstructured data such as texts through clustering analysis is paramount to many big data applications. However, document clustering is challenged by the computational complexity of the underlying methods and the high dimensionality of data, especially when the number of required clusters is large. A fine-grained clustering solution is required to understand a data set that represents heterogeneous topics such as social media data. This paper presents the *Fine-Grained document Clustering via Ranking* (FGCR) approach which leverages the search engine capability of handling big data efficiently. Ranking scores from a search engine are used to calculate dynamic clusters' representations called *loci* in an unsupervised learning setting. Clustering decisions are efficiently made based on an optimal selection from a small subset of loci instead of the entire cluster set as in the conventional centroid-based clustering. A comprehensive empirical study on several social media data sets shows that FGCR is able to produce insightful and accurate fine-grained solution. Moreover, it is magnitudes faster and requires less computational resources compared to other state-of-the-art document clustering approaches.

6 Conclusion and future work

We have introduced FGCR, a novel text clustering via ranking algorithm. FGCR creates a fine-grained clustering solution that is suitable to partition a large text collection with numerous topics such as social media data. The key benefit of FGCR is that its complexity is not affected by the total number of clusters in the data. Using loci and relevant clusters, FGCR does not need to scan all of the data to make clustering decisions. The extensive empirical study of FGCR using several social media data sets and comparing other clustering algorithms showed that the proposed method has a significantly higher clustering quality and requires minimal computational resources.

big data and cognitive computing

MDPI

Article

Fine-Grained Algorithm for Improving KNN Computational Performance on Clinical Trials Text Classification

Jasmir Jasmir [1,2,*], Siti Nurmaini [3] and Bambang Tutuko [3]

[1] Doctoral Program of Engineering Science, Faculty of Engineering, Universitas Sriwijaya, Palembang 30128, Indonesia
[2] Computer Engineering, Universitas Dinamika Bangsa, Jambi 36138, Indonesia
[3] Intelligent System Research Group, Universitas Sriwijaya, Palembang 30128, Indonesia; sitinurmaini@gmail.com (S.N.); bambang_tutuko@unsri.ac.id (B.T.)
* Correspondence: ijay_jasmir@yahoo.com

TELKOMNIKA Telecommunication, Computing, Electronics and Control
Vol. 19, No. 3, June 2021, pp. 886~892
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI: 10.12928/TELKOMNIKA.v19i3.18357

Bigram feature extraction and conditional random fields model to improve text classification clinical trial document

Jasmir Jasmir[1], Siti Nurmaini[2], Reza Firsandaya Malik[3], Bambang Tutuko[4]
[1]Faculty of Engineering, Universitas Sriwijaya, Palembang, Indonesia
[1]Computer Engineering, Universitas Dinamika Bangsa, Jambi, Indonesia
[2,4]Intelligence System Research Group, Universitas Sriwijaya, Palembang, Indonesia
[3]Communication Networks and Information Security Research Lab, Universitas Sriwijaya, Palembang, Indonesia

INTERNATIONAL CONFERENCE ICONISCSE INTERNATIONAL CONFERENCE ON INFORMATION SYSTEM COMPUTER SCIENCE AND ENGINEERING

The 2021 2nd International Conference on Information System, Computer Science and Engineering (ICONISCSE 2021)

January, 15th 2021

Letter of Acceptance (LoA)

Dear **Dr. Jasmir Jasmir**,

Congratulations - We are pleased to inform you that your paper:

#1570692454 : "Word Embedding for Eligibility Classification of Clinical Trials Document using Deep Learning"

has been **ACCEPTED** and will be presented in the ICONISCSE 2021 parallel session. The conference will be held on **March, 4th-5th 2021** in **Palembang Indonesia**.

Turkish Journal of Computer and Mathematics Education *Vol.12 No.14(2021), 2784- 2792*

Research Article

Improving Eligibility Classification on Clinical Trials Document using Bidirectional Long Short Term Memory Recurrent Neural Network

Jasmir Jasmir[a,b], Siti Nurmaini[c] , Reza Firsandaya Malik[d], Bambang Tutuko[c]

[a] Doctoral Program of Engineering Science, Faculty of Engineering Universitas Sriwijaya, Palembang Indonesia
[b]Computer Engineering, Universitas Dinamika Bangsa Jambi Indonesia
[c]Intelligence System Research GroupFaculty of Computer ScienceUniversitas Sriwijaya Palembang Indonesia
[c]Communication Networks and Information Security Research Lab, Faculty of Computer Science Universitas Sriwijaya, Palembang Indonesia

ATLANTIS PRESS

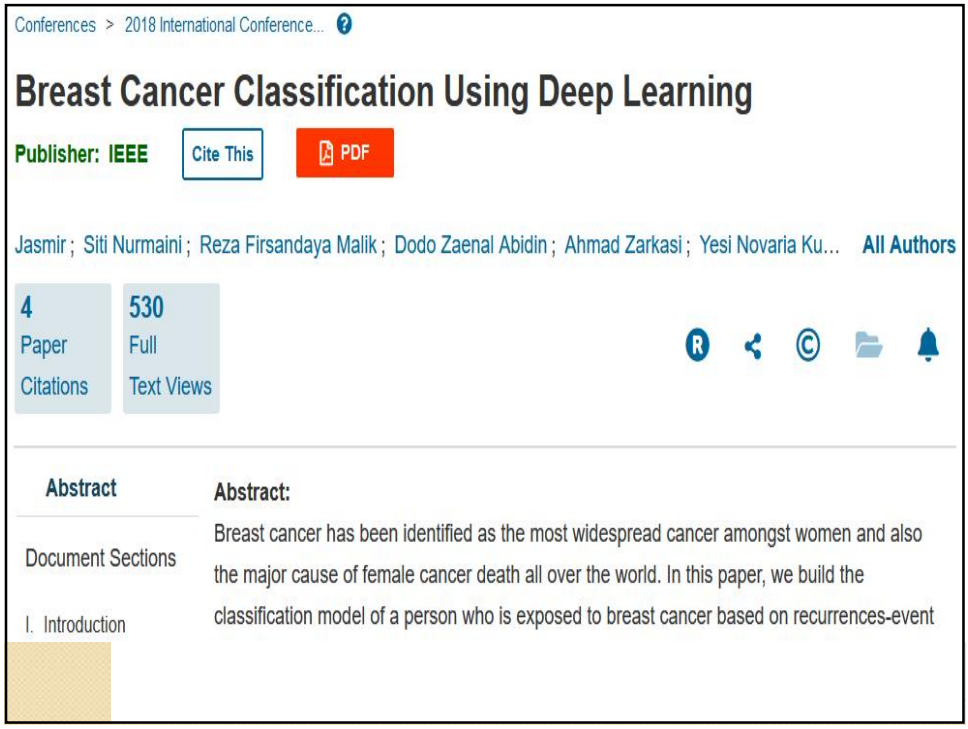
Advances in Intelligent Systems Research, volume 172

Sriwijaya International Conference on Information Technology and Its Applications (SICONIAN 2019)

Text Classification of Cancer Clinical Trials Documents Using Deep Neural Network and Fine Grained Document Clustering

Jasmir JASMIR[1,], Siti NURMAINI[2,] Reza Firsandaya MALIK[3*,] Dodo Zaenal ABIDIN[1]

[1].*Intelligence System Research Group, STIKOM Dinamika Bangsa Jambi Indonesia*
[2].*Intelligence System Research Group, Universitas Sriwijaya Palembang Indonesia*
[3].*Communication Networks and Information Security Research Lab, Universitas Sriwijaya, Palembang Indonesia*
Corresspondent author : rezafm@unsri.ac.id



Conferences > 2018 International Conference...

Breast Cancer Classification Using Deep Learning

Publisher: IEEE | Cite This | PDF

Jasmir ; Siti Nurmaini ; Reza Firsandaya Malik ; Dodo Zaenal Abidin ; Ahmad Zarkasi ; Yesi Novaria Ku... All Authors

4 Paper Citations | 530 Full Text Views

Abstract

Document Sections

I. Introduction

Abstract:

Breast cancer has been identified as the most widespread cancer amongst women and also the major cause of female cancer death all over the world. In this paper, we build the classification model of a person who is exposed to breast cancer based on recurrences-event

Accredited Ranking SINTA 2
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online on: http://jurnal.iaii.or.id

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 6 No. 4 (2022) 505 - 510 ISSN Media Electronic: 2580-0760

Bidirectional Long Short-Term Memory and Word Embedding Feature for Improvement Classification of Cancer Clinical Trial Document

Jasmir Jasmir[1], Willy Riyadi[2], Silvia Rianti Agustini[3], Yulia Arvita[4], Despita Meisak[5], Lies Aryani[6]
[1,2]Computer Engineering, Universitas Dinamika Bangsa Jambi Indonesia
[3,4,5,6]Information System, Universitas Dinamika Bangsa Jambi Indonesia
[1]ijay_jasmir@yahoo.com, [2]wriyadi5@gmail.com, [3]silviarianti7@gmail.com, [4]yulia_arvita@yahoo.co.id,
[5]liesaryani6@gmail.com, [6]despitam88@gmail.com

- https://www.sciencedirect.com/
- https://www.mdpi.com/
- https://www.hindawi.com/
- https://scholar.google.com/
- https://www.techscience.com/journals

Dalam basis data scopus dikenal dengan istilah quartil. Quartil merupakan pemeringkatan dari kumpulan jurnal dalam scope atau kategori yang sama. Quartil dibagi menjadi 4 yaitu quartil 1, quartil 2, quartil 3, dan quartil 4 atau bisa juga disebut Q1, Q2, Q3, dan Q4. Masing-masing jurnal dari basis data scopus ditentukan quartilnya dari nilai percentile. Perhatikan gambar berikut:

https://ilmubersama.com/2020/07/01/menghitung-percentile-dalam-penentuan-quartil-dari-basis-data-scopus/