AIP Conference Proceedings

scitation.org/journal/apc

Volume 2470

International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

Leveraging Smart Engineering

Surabaya, Indonesia • 25–26 August 2021

Editors • Joniarto Parung, Nemuel Daniel Pah, Putu Doddy Sutrisna and Marcellinus Ferdinand Suciadi

InCITE 2021

Leveraging Smart Engineering

Faculty of Engineering Universitas Surabaya

UBAYA UNIVERSITAS SURABAYA

AIP Publishing

InCITE 2021: VIRTUAL CONFERENCE OPEN AUDIENCE ATTENDANCE TO JOIN InCITE 2021
Registration Link *https://forms.gle/ZUjcCcFirshPBAfWA*

Program Book of InCITE 2021
Download Link https://bit.ly/program-book-of-incite-2021

Scientific Committee (International Reviewers)

InCITE 2021 receives very kind supports from world class scholars to be our international reviewers or international scientific committees. This will guarantee the presence of high quality papers from our reviewing process as the accepted papers will undergo rigorous review. Thus, the conference organizers would like to express their gratitude to the following scholars who will serve as the international scientific committees.

- Prof. Benny Tjahjono, Ph.D. (Coventry University, Coventry, England)
- Prof. Dr. Anton Satria Prabuwono (King Abdulaziz University, Jeddah, Saudi Arabia)
- Prof. Dr. Ing. Ir. I Made Londen Batan M.Eng. (Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia)
- Prof. Dr. Ir. Joko Lianto Buliali (Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia)
- Prof. Dr. Ir. Judy Retti B. Witono, M.App.Sc. (Universitas Katolik Parahyangan, Bandung, Indonesia)
- Prof. Dr. Ir. Wahyudi Sutopo, S.T., M.Si., IPM (Universitas Sebelas Maret, Solo, Indonesia)
- Prof. Ir. Arif Djunaidy, M.Sc., Ph.D. (Institut Teknologi Sepuluh Nopember (ITS), Surabaya, Indonesia)
- Prof. Ir. Djoko Budiyanto, M.Eng., Ph.D. (Universitas Atma Jaya, Yogyakarta, Indonesia)
- Prof. Ir. Joniarto Parung, M.M.B.A.T., Ph.D. (Universitas Surabaya, Surabaya, Indonesia)
- Prof. Ir. Lieke Riadi, Ph.D. (Universitas Surabaya, Surabaya, Indonesia)
- Prof. Itthisek Nilkhamhang, Ph.D. (Sirindhorn International Institute of Technology, Pathum Thani, Thailand)
- Prof. Dr. Pavel Albores (MIET, FHEA) (Aston University, Birmingham, England)
- Assoc. Prof. Bertha Maya Sopha, S.T., M.Sc., Ph.D. (Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia)
- Assoc. Prof. Dina Natalia Prayogo, S.T., M.Sc. (Universitas Surabaya, Surabaya, Indonesia)
- Assoc. Prof. Dr-Ing. Amalia Suzianti, S.T., M.Sc. (Universitas Indonesia (UI), Jakarta, Indonesia)
- Assoc. Prof. Dr. A. F. M. Saifuddin Saif (American International University, Bangladesh)
- Assoc. Prof. Dr. Andi Cakravastia Arisaputra Raja, S.T., M.T. (Institut Teknologi Bandung (ITB), Bandung, Indonesia)
- Assoc. Prof. Dr. Benjawan Srisura (Assumption University of Thailand, Bangkok, Thailand)
- Assoc. Prof. Dr. Diana Purwitasari, S.Kom., M.Sc. (Institut Teknologi Sepuluh Nopember (ITS), Surabaya)
- Assoc. Prof. Dr. Dra. Amelia Santoso, M.T. (Universitas Surabaya, Surabaya, Indonesia)
- Assoc. Prof. Dr. Dra. Ir. Evy Herowati, M.T. (Universitas Surabaya, Surabaya, Indonesia)
- Assoc. Prof. Dr. Emma Savitri, S.T., M.Sc. (Universitas Surabaya, Surabaya, Indonesia)
- Assoc. Prof. Dr. Ir. Anas Ma'ruf, M.T. (Institut Teknologi Bandung (ITB), Bandung, Indonesia)
- Assoc. Prof. Dr. Ir. Puguh Setyopratomo, M.T. (Universitas Surabaya, Surabaya, Indonesia)
- Assoc. Prof. Dr. Ir. Susila Candra, M.T. (Universitas Surabaya, Surabaya, Indonesia)
- Assoc. Prof. Dr. Joko Siswantoro, S.Si., M.Si. (Universitas Surabaya, Surabaya, Indonesia)
- Assoc. Prof. Dr. Ludovic F. Dumee (Khalifa University, Uni Arab Emirates (UAE))
- Assoc. Prof. Dr. Mohd Sanusi Azmi (Universiti Teknikal Malaysia Melaka, Melaka)
- Assoc. Prof. Dr. Ridi Ferdiana, S.T., M.T. (Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia)
- Assoc. Prof. Dr. Rosita Meitha Surjani, S.T., M.T. (Universitas Surabaya, Surabaya, Indonesia)
- Assoc. Prof. Dr.Eng. Wahyudiono (Nagoya University, Nagoya, Japan)
- Assoc. Prof. Elieser Tarigan, S.Si., M.Eng., Ph.D. (Universitas Surabaya, Surabaya, Indonesia)
- Assoc. Prof. Ir. Eric Wibisono, Ph.D., IPU. (Universitas Surabaya, Surabaya, Indonesia)
- Assoc. Prof. Ir. Hanung Adi Nugroho, Ph.D. (MIEEE, MACM) (Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia)
- Assoc. Prof. Restu Kartiko Widi, S.Si., M.Si., Ph.D. (Universitas Surabaya, Surabaya, Indonesia)
- Assoc. Prof. Satvasheel Powar, Ph.D. (Dalarna University, Sweden)
- Assoc. Prof. Tan Kay Chuan, Ph.D. (National University of Singapore (NUS), Singapore)
- Assoc. Prof. Teguh Bharata Adji, S.T., M.T., M.Eng., Ph.D. (Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia)
- Assoc. Prof. The, Jaya Suteja, S.T., M.Sc., Ph.D. (Universitas Surabaya, Surabaya, Indonesia)
- Assoc. Prof. Timotius Pasang, Ph.D. (Oregon Institute of Technology, USA)
- Alexander Yohan, S.Kom., M.I.M., Ph.D. (Universitas Surabaya, Surabaya, Indonesia)
- Djuwari, S.T., Ph.D. (Universitas Surabaya, Surabaya, Indonesia)
- Dr. Awang Hendrianto Pratomo (UPN "Veteran" Yogyakarta, Indonesia)
- Dr. Delta Ardy Prima, S.ST., M.T. (Universitas Surabaya, Surabaya, Indonesia)
- Dr. Drs. Muhammad Rosiawan, M.T. (Universitas Surabaya, Surabaya, Indonesia)
- Dr. Fredy Purnomo, S.Kom, M.Kom. (Universitas Bina Nusantara, Jakarta, Indonesia)
- Dr. Hazrul Iswadi, S.Si., M.Si. (Universitas Surabaya, Surabaya, Indonesia)
- Dr. Hendri Himawan Triharminto (Akademi Angkatan Udara, Yogyakarta, Indonesia)
- Dr. Indri Hapsari, S.T., M.T. (Universitas Surabaya, Surabaya, Indonesia)
- Dr. Jimmy, S.T., M.I.S. (Universitas Surabaya, Surabaya, Indonesia)
- Dr. Khoiruddin, S.T., M.T. (Institut Teknologi Bandung (ITB), Bandung, Indonesia)
- Dr. Mohamad Rezi bin Abdul Hamid (University Putra Malaysia, Seri Kembangan, Malaysia)
- Dr. Naniek Utami Handayani, S.Si., M.T. (Universitas Diponegoro, Semarang, Indonesia)
- Dr. Yenny Sari, S.T., M.Sc. (Universitas Surabaya, Surabaya, Indonesia)
- Dr.rer.nat. Lanny Sapei, S.T., M.Sc. (Universitas Surabaya, Indonesia)
- Firman Kurniawansyah, S.T., M.Eng.Sc., Ph.D. (Institut Teknologi Sepuluh Nopember (ITS), Surabaya)
- Nemuel Daniel Pah, S.T., M.Eng., Ph.D. (Universitas Surabaya, Surabaya, Indonesia)
- Ratna Surya Alwi, S.T., M.Si. Ph.D. (Universitas Fajar, Makassar, Indonesia)
- Drs. Heru Arwoko, M.T. (Universitas Surabaya, Surabaya, Indonesia)
- Hendi Wicaksono Agung Darminto, S.T., M.T. (Universitas Surabaya, Surabaya, Indonesia)

- Liliana, S.T., M.MSI. (Universitas Surabaya, Surabaya, Indonesia)
- Maya Hilda Lestari Louk, S.T., M.Sc. (Universitas Surabaya, Surabaya, Indonesia)
- Mohammad Farid Naufal, S.Kom., M.Kom. (Universitas Surabaya, Surabaya, Indonesia)
- Monica Widiasri, S.Kom., M.Kom. (Universitas Surabaya, Surabaya, Indonesia)
- Vincentius Riandaru Prasetyo, S.Kom., M.Cs. (Universitas Surabaya, Surabaya, Indonesia)
- Yuana Elly Agustin, S.T., M.Sc. (Universitas Surabaya, Surabaya, Indonesia)
- Yunus Fransiscus, S.T., M.Sc. (Universitas Surabaya, Surabaya, Indonesia)
- Argo Hadi Kusumo, S.T., M.B.A. (Universitas Surabaya, Surabaya, Indonesia)
- I Made Ronyastra, S.T., M.T. (Universitas Surabaya, Surabaya, Indonesia)

Steering Committee

Chair	: Asst. Prof. Putu Doddy Sutrisna, S.T., M.Sc., Ph.D.
Honorary Members	: Prof. Ir. Joniarto Parung, M.M.B.A.T., Ph.D. Prof. Ir. Lieke Riadi, Ph.D. Assoc. Prof. Ir. Eric Wibisono, Ph.D., IPU Assoc. Prof. Ir. Markus Hartono, S.T., M.Sc., Ph.D. Assoc. Prof. Djuwari, S.T., Ph.D. Assoc. Prof. Dr. Dra. Amelia Santoso, M.T. Asst. Prof. Nemuel Daniel Pah, S.T., M.Eng., Ph.D. Assoc. Prof. Dr. Emma Savitri, S.T., M.Sc. Assoc. Prof. Elieser Tarigan, S.Si., M.Eng., Ph.D.
Members	: Assoc. Prof. Jaya Suteja, S.T., M.Sc., Ph.D. Assoc. Prof. Ir. Hudiyo Firmanto, M.Sc., Ph.D. Assoc. Prof. Restu Kartiko Widi, S.Si., M.Si., Ph.D. Assoc. Prof. Dr. Joko Siswantoro, S.Si., M.Si. Assoc. Prof. Dr. Indri Hapsari, S.T., M.T. Asst. Prof. Susilo Wibowo, S.T., M.Eng. Asst. Prof. Arum Soesanti, S.T., M.T.

Organising Committee

Chair	: Asst. Prof. Dr. Jimmy, S.T., M.I.S
Vice Chair	: Asst. Prof. Yuwono Budi Pratiknyo, S.T., M.T.
Secretariat	: Ms. Liliana, S.T., M.MSI. Mr. Marcellinus Ferdinand Suciadi, S.T., M.Comp.
Treasurer	: Ms. Rafina Destiarti Ainul, S.ST., MT.
Program	: Ms. Maya Hilda Lestari Louk, S.T., M.Sc. Mr. Herman Susanto, S.T.,M.Sc.
Website	: Mr. Daniel Soesanto, S.T., M.M. Mr. Felix Handani, S.Kom., M.Kom.
Design	: Ms. Tryza Adelia, S.Sn., M.Inf.Tech.

MENU SIGN IN/REGISTER

AIP Publishing Conference Proceedings

HOME BROWSE MORE

Table of Contents

INTERNATIONAL CONFERENCE ON INFORMATICS, TECHNOLOGY, AND ENGINEERING 2021 (InCITE 2021): Leveraging Smart Engineering

‹ PREV NEXT ›

Conference date: 25–26 August 2021
Location: Surabaya, Indonesia
ISBN: 978-0-7354-4180-4
Editors: Joniarto Parung, Nemuel Daniel Pah, Putu Doddy Sutrisna and Marcellinus Ferdinand Suciadi Suciadi
Volume number: 2470
Published: Apr 25, 2022

DISPLAY : 20 50 100 all

PRELIMINARY

BROWSE VOLUMES

No Access . April 2022

Preface: International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

more...

AIP Conference Proceedings **2470**, 010001 (2022); https://doi.org/10.1063/12.0009021

No Access . April 2022

Committees: International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

more...

AIP Conference Proceedings **2470**, 010002 (2022); https://doi.org/10.1063/12.0009708

SMART DESIGN AND INNOVATION

No Access . April 2022

How energy efficiency, smart factory, and mass personalization affect companies in industry 4.0

Yu-Chung Tsao, Felix Arril Simbara Barus and Chien-Wei Ho

AIP Conference Proceedings **2470**, 020001 (2022); https://doi.org/10.1063/5.0080526

SHOW ABSTRACT

No Access . April 2022

Kansei engineering and product-service systems (KEPSS) integrative framework for customer-centered experience

Markus Hartono

AIP Conference Proceedings **2470**, 020002 (2022); https://doi.org/10.1063/5.0080177

SHOW ABSTRACT

No Access . April 2022

The role of supply chain performance factors to determine the firm performance

Yoni Kristiawan, Edi Purwanto and Rustono Farady Marta

AIP Conference Proceedings **2470**, 020003 (2022); https://doi.org/10.1063/5.0080235

SHOW ABSTRACT

BROWSE VOLUMES

No Access . April 2022

Real-time monitoring design for make-to-order industry

Senapati C. Adisasmito, Pradipta Deffinika Pamungkas and Anas Ma'ruf

AIP Conference Proceedings **2470**, 020004 (2022); https://doi.org/10.1063/5.0080747

SHOW ABSTRACT

No Access . April 2022

Reducing production process lead time using Value Stream Mapping and Kaizen approaches: A case study in the musical instrument industry

Indra Setiawan and Hernadewita

AIP Conference Proceedings **2470**, 020005 (2022); https://doi.org/10.1063/5.0080159

SHOW ABSTRACT

No Access . April 2022

Simple heuristics for scheduling apache airflow: A case study at PT. X

Hans Tanto, Indriati N. Bisono and Hanijanto Soewandi

AIP Conference Proceedings **2470**, 020006 (2022); https://doi.org/10.1063/5.0081042

SHOW ABSTRACT

BROWSE VOLUMES

Quality improvement with PDCA approach and design of experiment method in single socks industry in Indonesia

Hibarkah Kurnia, Choesnul Jaqin and Humiras Hardi Purba

AIP Conference Proceedings **2470**, 020007 (2022); https://doi.org/10.1063/5.0080179

SHOW ABSTRACT

No Access . April 2022

Integration of corporate social responsibility for improved company performance: Evidence from the Indonesian manufacturing industry

Esti Dwi Rinawiyanti, Xueli Huang and Sharif As-Saber

AIP Conference Proceedings **2470**, 020008 (2022); https://doi.org/10.1063/5.0080731

SHOW ABSTRACT

No Access . April 2022

Community mobility during Covid-19 pandemic and tourism performance: Data mining approach

Gunawan

AIP Conference Proceedings **2470**, 020009 (2022); https://doi.org/10.1063/5.0080170

SHOW ABSTRACT

BROWSE VOLUMES

Inventory system improvement for poultry

Noverta Brilly Leksana Putra, Indri Hapsari and Dina Natalia Prayogo

AIP Conference Proceedings **2470**, 020010 (2022); https://doi.org/10.1063/5.0080160

SHOW ABSTRACT

No Access . April 2022

Users' perception of digital prototypes in Indonesian fashion industry: A qualitative study

Christabel Parung and Prayogo Waluyo

AIP Conference Proceedings **2470**, 020011 (2022); https://doi.org/10.1063/5.0080178

SHOW ABSTRACT

SMART MANUFACTURING AND PROCESSES

No Access . April 2022

Dynamic study of batch milk cooling process at KUD SAE Pujon

Rudy Agustriyanto, Puguh Setyopratomo and Endang Srihari

AIP Conference Proceedings **2470**, 030001 (2022); https://doi.org/10.1063/5.0080176

SHOW ABSTRACT

BROWSE VOLUMES

Recent advances and application of Selective Laser Melting (SLM) technology in the aerospace industry

Anel Yerubayeva, Essam Shehab and Md. Hazrat Ali

AIP Conference Proceedings **2470**, 030002 (2022); https://doi.org/10.1063/5.0080173

SHOW ABSTRACT

POWER SYSTEM AND SMART ENERGY MANAGEMENT

No Access . April 2022

Immobilization of xylanase on acid pretreatment bentonite as green biocatalyst

Lieke Riadi, Ruth Chrisnasari, Joshua Kristanto, Cahaya Caesar Bigravida and Meyta Sanoe

AIP Conference Proceedings **2470**, 040001 (2022); https://doi.org/10.1063/5.0080318

SHOW ABSTRACT

No Access . April 2022

Rheological behavior and antioxidant activity of carrageenan extracted from Green seaweed (*Eucheuma cottonii*) using alkaline solution at low temperature

Puguh Setyopratomo and Lanny Sapei

AIP Conference Proceedings **2470**, 040002 (2022); https://doi.org/10.1063/5.0080286

SHOW ABSTRACT

BROWSE VOLUMES

No Access . April 2022

Simulation of impact azimuth angle on specific energy output of a fixed mounting rooftop PV system in Jakarta, Indonesia

Elieser Tarigan

AIP Conference Proceedings **2470**, 040003 (2022); https://doi.org/10.1063/5.0080228

SHOW ABSTRACT

No Access . April 2022

Rice husk ash for the stabilization of the outer interfacial layer of W/O/W double emulsion

Lanny Sapei, Rudy Agustriyanto, Endang Wahyu Fitriani, Zerravym Levy and Cindy Sumampouw

AIP Conference Proceedings **2470**, 040004 (2022); https://doi.org/10.1063/5.0080285

SHOW ABSTRACT

No Access . April 2022

The effects of material to solvent ratio on the performances of natural dyes extraction

Putu Doddy Sutrisna, Meyta Sanoe, Rifando Gogo Adiyaksa, Kristina Wahyu Agustine, Hadiyatni Rita Priyantini, Prayogo Widyastoto Waluyo and I. Made Ronyastra

AIP Conference Proceedings **2470**, 040005 (2022); https://doi.org/10.1063/5.0080171

SHOW ABSTRACT

BROWSE VOLUMES

BROWSE VOLUMES

One-step electrochemical synthesis of silica-coated magnetite nanofluids

Delyana Ratnasari, W. Widiyastuti and Heru Setyawan

AIP Conference Proceedings **2470**, 040009 (2022); https://doi.org/10.1063/5.0080982

SHOW ABSTRACT

No Access . April 2022

Synthesis and characterization of chitosan-Allium sativum film

Emma Savitri, Maria Irine Tjahayani and Anastasia Apriyani Ngene Say

AIP Conference Proceedings **2470**, 040010 (2022); https://doi.org/10.1063/5.0080203

SHOW ABSTRACT

No Access . April 2022

The effect of sodium lauryl sulfate on silica nanofluid stabilization using microbubble method

Ratri Sekaringgalih, I. Made Joni, W. Widiyastuti and Heru Setyawan

AIP Conference Proceedings **2470**, 040011 (2022); https://doi.org/10.1063/5.0080985

SHOW ABSTRACT

No Access . April 2022

Textile wastewater purification using corona discharge method

BROWSE VOLUMES

AIP Conference Proceedings **2470**, 040012 (2022); https://doi.org/10.1063/5.0080740

SHOW ABSTRACT

THE ROLE OF IT IN INNOVATION ENHANCEMENT

No Access . April 2022

Hate speech content detection system on Twitter using K-nearest neighbor method

Vincentius Riandaru Prasetyo and Anton Hendrik Samudra

AIP Conference Proceedings **2470**, 050001 (2022); https://doi.org/10.1063/5.0080185

SHOW ABSTRACT

No Access . April 2022

School finder, intelligent recommendation system for elementary school selection

Susana Limanto, Endah Asmawati and Yusuf Wira Kencana Putra

AIP Conference Proceedings **2470**, 050002 (2022); https://doi.org/10.1063/5.0080461

SHOW ABSTRACT

No Access . April 2022

BROWSE VOLUMES

decision tree algorithm in data mining with KNIME (for recommending the topic of undergraduate's thesis)

Yenny Sari, Vincentius Riandaru Prasetyo and Kevin Liyansah

AIP Conference Proceedings **2470**, 050003 (2022); https://doi.org/10.1063/5.0081214

SHOW ABSTRACT

No Access . April 2022

Weather image classification using convolutional neural network with transfer learning

Mohammad Farid Naufal and Selvia Ferdiana Kusuma

AIP Conference Proceedings **2470**, 050004 (2022); https://doi.org/10.1063/5.0080195

SHOW ABSTRACT

No Access . April 2022

Student performance prediction in higher education: A comprehensive review

Ellysa Tjandra, Sri Suning Kusumawardani and Ridi Ferdiana

AIP Conference Proceedings **2470**, 050005 (2022); https://doi.org/10.1063/5.0080187

SHOW ABSTRACT

No Access . April 2022

BROWSE VOLUMES

conduct: An empirical evaluation of methods

Jimmy and Vincentius Riandaru Prasetyo

AIP Conference Proceedings **2470**, 050006 (2022); https://doi.org/10.1063/5.0080182

SHOW ABSTRACT

No Access . April 2022

Implementing directed pairwise judgement approach in web-based AHP survey application to reduce inconsistency ratio

Daniel Hary Prasetyo

AIP Conference Proceedings **2470**, 050007 (2022); https://doi.org/10.1063/5.0080169

SHOW ABSTRACT

No Access . April 2022

Leaf geometric properties measurement using computer vision system based on camera parameters

Joko Siswantoro, Ida Bagus Made Artadana and Muhammad Z. F. N. Siswantoro

AIP Conference Proceedings **2470**, 050008 (2022); https://doi.org/10.1063/5.0080190

SHOW ABSTRACT

No Access . April 2022

A design of secure supply chain management system with

BROWSE VOLUMES

Alexander Yohan, Nai-Wei Lo and Kevin Valentino

AIP Conference Proceedings **2470**, 050009 (2022); https://doi.org/10.1063/5.0080266

SHOW ABSTRACT

No Access . April 2022

Implementation of behavior tree for creating an in-game cut-scene

Delta Ardy Prima

AIP Conference Proceedings **2470**, 050010 (2022); https://doi.org/10.1063/5.0080233

SHOW ABSTRACT

No Access . April 2022

The online attendance system models for educational institutions

Andre and Marcellinus Ferdinand Suciadi

AIP Conference Proceedings **2470**, 050011 (2022); https://doi.org/10.1063/5.0080180

SHOW ABSTRACT

No Access . April 2022

Genetic algorithm with adaptive diversification and intensification for the vehicle routing problem

BROWSE VOLUMES

AIP Conference Proceedings **2470**, 050012 (2022); https://doi.org/10.1063/5.0080197

SHOW ABSTRACT

Resources

AUTHOR

LIBRARIAN

ADVERTISER

General Information

ABOUT

CONTACT

HELP

PRIVACY POLICY

TERMS OF USE

FOLLOW AIP PUBLISHING:

BROWSE VOLUMES

Scitation

specified within the article.

BROWSE VOLUMES

Preface: International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

Cite as: AIP Conference Proceedings **2470**, 010001 (2022); https://doi.org/10.1063/12.0009021
Published Online: 25 April 2022

Preface: International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

Welcome to the 3rd biannual International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021). This year, InCITE is organized from Surabaya, Indonesia and researchers from around the world discuss current challenges and technological solutions via an online platform on August 25-26, 2021.

InCITE 2021 presents "Leveraging Smart Engineering" as its central theme. In response to several world challenges, such as sustainable development, global convergence of information and communications technologies, along with smart systems as opportunities as well as challenges in developments for better industries, it is considered important to discover innovative approaches from science and engineering perspectives. Innovation suggests the introduction of novelty to create better solutions. Innovation in engineering and science requires contributions from multidisciplinary sectors, involving industries, practitioners, researchers, and academics.

International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)
AIP Conf. Proc. 2470, 010001-1–010001-1; https://doi.org/10.1063/12.0009021
Published by AIP Publishing. 978-0-7354-4180-4/$30.00

Committees: International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)

Cite as: AIP Conference Proceedings **2470**, 010002 (2022); https://doi.org/10.1063/12.0009708
Published Online: 25 April 2022

WELCOME FROM InCITE 2021 STEERING COMMITTEE

It is a great pleasure to welcome all of you to the 3rd Bi-Annual International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021) held by the Faculty of Engineering, University of Surabaya (UBAYA). The first and second InCITE have been successfully held in Bali, Indonesia in 2017 and 2019. Hence, now we are delighted to host the third InCITE through online media due to the Covid-19 pandemic situation.

There are 37 papers that have been selected to be presented in InCITE 2021. The papers were written by experts not only from Indonesia, but also from different parts of the world. The main theme of this conference is Leveraging Smart Engineering in response to the current and future Industrial Revolution 4.0 that should be handled by every country in the world. We hope through this conference, all participants will be able to know each other and interact to develop future collaboration.

We would like to express our sincere gratitude to the Keynote speakers, International Scientific Committee, Steering Committee, and Organizing Committee for their huge efforts to make this conference successful.

Thank you all for your support and attendance at InCITE 2021. Please enjoy the conference!

Asst. Prof. Doddy Sutrisna, Ph.D.
Steering Committee

International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)
AIP Conf. Proc. 2470, 010002-1–010002-2; https://doi.org/10.1063/12.0009708
Published by AIP Publishing. 978-0-7354-4180-4/$30.00

WELCOME FROM InCITE 2021 ORGANIZING COMMITTEE

Welcome to InCITE 2021! The third bi-annual international conference on engineering domain conducted by the Faculty of Engineering, The University of Surabaya (UBAYA). Due to the COVID-19 pandemic, InCITE 2021 is held as an online conference. Online conference opens the opportunity for many researchers around the globe to share their findings and learn from other global researchers with less restrictions.

InCITE 2021 invites three keynote speakers, well reputable global researchers in their research domain from Australia and Taiwan. Following each keynote session are two presentation sessions run in parallel.

This year, we received 66 papers submitted by researchers from four distinct countries (i.e., first author's country of origin): Indonesia, Australia, Taiwan, and Kazakhstan.

We employed a double-blind review to ensure a high standard and a minimum level of bias in the reviewing processes. This resulted in 56% of the submissions were accepted and will be published to the AIP Conference Proceedings.

Authors of all accepted papers are to disseminate their findings during InCITE 2021 conference between 25 to 26 of August 2021. This presents a great opportunity for everyone, including the researchers, to discuss and further improve current achievements.

We thank all keynote speakers, presenters, and reviewers/scientific committees for the generous supports. We thank the University of Surabaya, the Faculty of Engineering UBAYA, and all InCITE 2021 committees that enable InCITE 2021.

We wish you a very pleasant and rich conference experience in InCITE 2021 and look forward to seeing you again on InCITE 2023! Thank you.

Yours sincerely,
Asst. Prof. Dr. Jimmy
InCITE 2021 Organizing Committee

Scopus Preview

Author Search Sources Create account Sign in

This author profile is generated by Scopus Learn more

Prasetyo, Vincentius Riandaru

Universitas Surabaya, Surabaya, Indonesia

https://orcid.org/0000-0003-1376-7270 Is this you? Connect to Mendeley account

Edit profile Set alert Potential author matches Export to SciVal

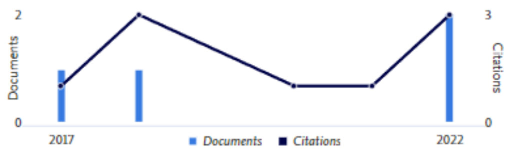
Metrics overview

4
Documents by author

9
Citations by 8 documents

2
h-index: View *h*-graph

Document & citation trends



Most contributed Topics 2016–2020

Information Retrieval; Question Answering System; Word List
2 documents

View all Topics

4 Documents Cited by 8 Documents 0 Preprints 3 Co-Authors 1 Topic 0 Awarded Grants Beta

Note:
Scopus Preview users can only view an author's last 10 documents, while most other features are disabled. Do you have access through your institution? Check your institution's access to view all documents and features.

Export all Add all to list Sort by Date (newest)

View list in search results format

View references

Set document alert

Conference Paper • *Open access*

Hate speech content detection system on Twitter using K-nearest neighbor method

Prasetyo, V.R., Samudra, A.H.

AIP Conference Proceedings, 2022, 2470, 050001

0
Citations

Hate speech content detection system on Twitter using K-nearest neighbor method

Cite as: AIP Conference Proceedings **2470**, 050001 (2022); https://doi.org/10.1063/5.0080185
Published Online: 25 April 2022

Vincentius Riandaru Prasetyo and Anton Hendrik Samudra

Hate Speech Content Detection System on Twitter using K-Nearest Neighbor Method

Vincentius Riandaru Prasetyo[1, a)] and Anton Hendrik Samudra[2, b)]

[1]*Department of Informatics, Faculty of Engineering, University of Surabaya, Surabaya 60293, Indonesia*
[2]*Faculty of Law, University of Surabaya, Surabaya 60293, Indonesia*

[a)] Corresponding author: vincent@staff.ubaya.ac.id
[b)] antonhendrik@staff.ubaya.ac.id

Abstract. Twitter is a social media platform that many Indonesians use to express their thoughts on a variety of topics. In Indonesia, the use of social media is governed by a law known as Information and Electronic Transactions Law. However, until now, the implementation of this law has been subpar. This is because there are still violations occurring, and no legal action has been taken against these violations. Hate speech is a common violation on Twitter. The goal of this research is to create a system that can detect potential violations of content on Twitter, particularly content containing hate speech. The k-nearest neighbor (KNN) method was used in this research, along with the feature extraction method TF-IDF. The system built will detect whether the tweet you want to post violates a specific article in the Information and Electronic Transactions Law. Based on model validation, model classifier built has accuracy value is 67.86%, with *K* value in the KNN method is 10. Meanwhile, based on user validation, the system created has an accuracy of 77%.

Keywords: Hate Speech Detection, Twitter, K-Nearest Neighbor, ITE Law

INTRODUCTION

Social media is a type of online communication medium that allows people to share various types of information such as text, images, sound, and video. Twitter is one of the most popular social media platforms. Twitter is a social media platform that allows users to share information in the form of “Tweets”. According to data from the Ministry of Communication and Information Technology, Indonesia is the world's fifth-largest Twitter user country, trailing only the United Kingdom [1]. The number of Twitter users in Indonesia has risen to 19.5 million and is expected to rise further.

The Information and Electronic Transactions Law governs the use of social media as a medium of electronic communication in Indonesia. The regulation is a law that governs the use and utilization of electronic technology to support activities and ensures that activities are carried out in accordance with the Indonesian people's social and cultural conditions. However, the regulation's implementation has been ineffective, as violations are still being reported on social media. Hate speech was one of the violations that occurred. Hate speech is defined as an act of communication carried out by an individual or group in the form of provocation, incitement, or insults directed at other individuals or groups based on factors such as race, color, ethnicity, gender, sexual orientation, nationality, religion, etc [2]. Some hate speech violations have resulted in appropriate legal action, such as the examples of hate speech via Twitter shown in Fig. 1. The tweet has been determined with decision number: 1521/Pid.Sus/2017/PN Jkt.Pst and appeal decision number 1940 K/Pid.Sus/2018, where the defendant was sentenced to two years in prison. The articles that can be used to prosecute hate speech offenders are: Article 156 of the Criminal Code, Article 156a of the Criminal Code, and Article 45A paragraph (2) jo. Article 28 paragraph (2) of the ITE Law. These three articles are the focus of this research.

International Conference on Informatics, Technology, and Engineering 2021 (InCITE 2021)
AIP Conf. Proc. 2470, 050001-1–050001-10; https://doi.org/10.1063/5.0080185
Published by AIP Publishing. 978-0-7354-4180-4/$30.00

FIGURE 1. Hate Speech Tweet That Has Received A Legal Ruling

However, there are many hate-speech tweets that have yet to be investigated by the police. This is due to the large number of hate speech tweets in circulation, which makes it difficult for police to track down the original source or provocateurs of hate speech. Figure 2 depicts an example of a tweet that did not result in legal action by the police. Another factor that contributes to the widespread spread of hate speech is social media users' ignorance of the consequences of the hate speech they post. The system developed in this research can assist Twitter users in first checking the tweet to be posted, whether it has the potential to violate certain legal articles, and the consequences if the tweet is still posted. The method used in this study is k-nearest neighbor.

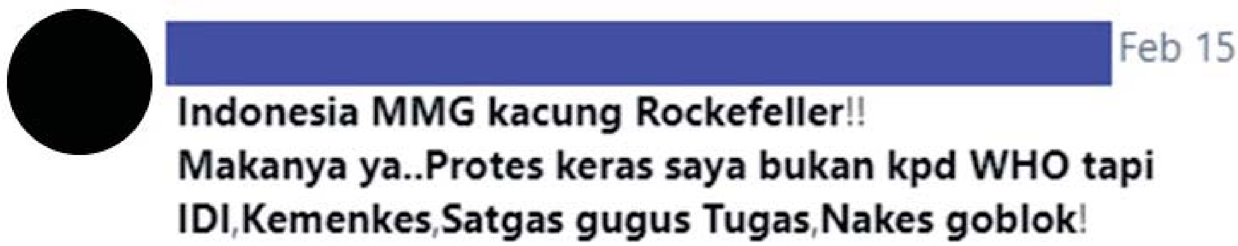


FIGURE 2. Hate Speech Tweet That Have Not Received Legal Action.

The k-nearest neighbor (KNN) is an instance-based learning method that stores all available data points (examples) and classifies new data points using a similarity measure to classify them. The KNN method works by assigning new unclassified examples to the class that contains the majority of its k-nearest neighbors. When the number of samples in the training dataset is large, this algorithm proves to be very effective in terms of reducing misclassification error. Another advantage of the KNN method over many other supervised learning methods such as support vector machine (SVM), decision tree, neural network, and so on is that it can easily deal with problems with class sizes of three or more [3].

RELATED WORKS

There has previously been research on the topic of detecting hate speech and using Twitter data. In their study, Alfina et al. [4] create a new dataset that includes hate speech in general, including hatred for religion, race, ethnicity, and gender. They also conducted a preliminary study utilizing a machine learning approach. They compared the performance of various features and machine learning algorithms for detecting hate speech. The extracted features were word n-grams with n=1 and n=2, character n-grams with n=3 and n=4, and negative sentiment. The classification was done with Nave Bayes, Support Vector Machine, Bayesian Logistic Regression, and Random Forest Decision Tree.

On the other hand, research conducted by Febriana and Budiarto [5] presents the process of developing a dataset that can be used to build a hate speech detection model from more than 1 million tweets successfully collected via Twitter API. Machine learning was used to perform basic preprocessing and preliminary research. The Latent Dirichlet Allocation (LDA) algorithm was used to extract the topic for each tweet in order to determine whether these topics are related to debate themes. The dataset was also subjected to pre-trained sentiment analysis, which resulted in polarity scores for each tweet. The number of positive and negative tweets is almost equal among the 83,752 tweets included in the analysis step.

Patihullah and Winarko [6] in their research compares the performance of Gated Recurrent Units (GRU) to that of other supervised methods such as Support Vector Machine, Naive Bayes, Random Forest, and Logistic Regression. Word2vec was used for feature extraction in their study because it has the ability to learn semantics between words. However, compared to TF and TF-IDF features, the use of word2vec in comparison methods yields lower accuracy. In contrast to previous research, Taradhita and Putra [7] propose a convolutional neural network method for classifying hate speech in Indonesian tweets. Twitter was used to collect datasets for both the training and testing stages. The collected tweets were divided into two categories: hate speech and non-hate speech. For feature extraction, they used the TF-IDF term weighting method.

This study employs a different classification method than previous studies, namely the KNN method. The KNN method was chosen because it is well-suited to problems with class sizes of three or more. Furthermore, if the number of datasets is increased, the KNN method can reduce classification errors [3]. This study employs the TF-IDF method for feature extraction because, according to the research conducted by Patihullah and Winarko [6], this method produces good accuracy. In addition, this research can specifically tell the user whether a tweet is included in hate speech or not and which rules are violated, so that the user knows the consequences if the tweet is still posted. This is not found in the studies that have been discussed previously.

RESEARCH METHODS

Collecting and Labeling Data

The first step in this study was to collect data in the form of hate speech-related tweets. The tweet data collected consists of Indonesian tweets. Crawling techniques are used in this data collection process. Crawling is a technique for retrieving the components of a web page and extracting the components contained within them. Obtaining a web URL is the first step in the crawling process. The URL is then downloaded in order to obtain a collection of pages from that URL. The obtained page is then extracted from the components that contain information related to the required data [8]. Crawling was carried out in this study using the Twitter API. The Twitter API is a developer-provided program that can be used to access information from Twitter. The Twitter API is commonly used for education, research, and analysis [9].

Following the collection of tweet data, the data must be labeled. Each tweet will be labeled with one of four labels: *"Pasal 156 KUHP", "Pasal 156a KUHP", "Pasal 45A ayat (2) jo. Pasal 28 ayat (2) UU ITE*", indicating that it violates certain articles, or "*Tidak Melanggar*", where the tweet does not violate the law. Several research assistants labeled the data after receiving guidance from ITE law experts.

Preprocessing data

Preprocessing occurs after the data has been collected and labeled. Preprocessing aims to obtain clean data in order to improve the accuracy of the detection process [10]. The first step in preprocessing is to remove symbols that frequently appear in tweet data, such as #, @, emoticons, and website links. Because the symbols have little influence on the detection process, this is done. The following stage is case folding, which converts all letters in a word to lowercase. This is done to avoid the use of words with the same meaning, such as "*SaYa*" and "*saya*".

Stemming is the next step in preprocessing. This is a process for determining the root of a word that will be implemented using the Nazief Andriani algorithm, which will be used in this study with the *Sastrawi 1.0.1* library. The Nazief Andriani algorithm was developed based on the broad Indonesian morphological rules, which were collected into one group and encapsulated into allowed affixes and disallowed affixes. This algorithm uses a basic word dictionary and supports recording, namely the rearrangement of words that have an excessive stemming process [11]. Wahyudi, et al [11] in their research also concluded that the Nazief Andriani algorithm gave better results than the Porter algorithm to support information retrieval systems. The next step after stemming is stop word removal. This step was used to eliminate words that had no effect on the categorization process, such as: *yang, dan, atau, ke, dari*, and so on. Tokenization is the final step in the preprocessing process. Tokenization is the process of dividing a sentence into words. Tokenization results will be saved as clean data.

Feature Extraction

The feature extraction process is carried out after the training data has been preprocessed. The Term Frequency - Inverse Document Frequency (TF-IDF) method can be used to extract features from text data. The TF-IDF method is a combination of weighting methods used to determine the importance of a word in a document. Formula (1) can be used to calculate TF-IDF values [12].

$$tf.idf = tf \times \log \frac{N}{df} \tag{1}$$

where *tf* is the term frequency, which is the frequency with which a word or term appears in a given document. Meanwhile, *idf* stands for inverted document frequency, which is the base 10 log value of the number of *N* documents divided by the *df* value. The *df* stands for document frequency, which is the number of documents that contain a specific term.

Training Data

After extracting the features, the training process begins. The goal of this process is to train the feature data using a specific method in order to create a classification model that can later be used to predict hate speech tweets. The KNN was used to train the data in this study. KNN is a classification method that computes the shortest distance between training and test data. The accuracy of this method is determined by the features it possesses. The accuracy of this method decreases as the number of relevant features decreases. This method is appropriate for training data that is large and noisy [13]. The stages of data classification by KNN method is shown in Fig. 3.

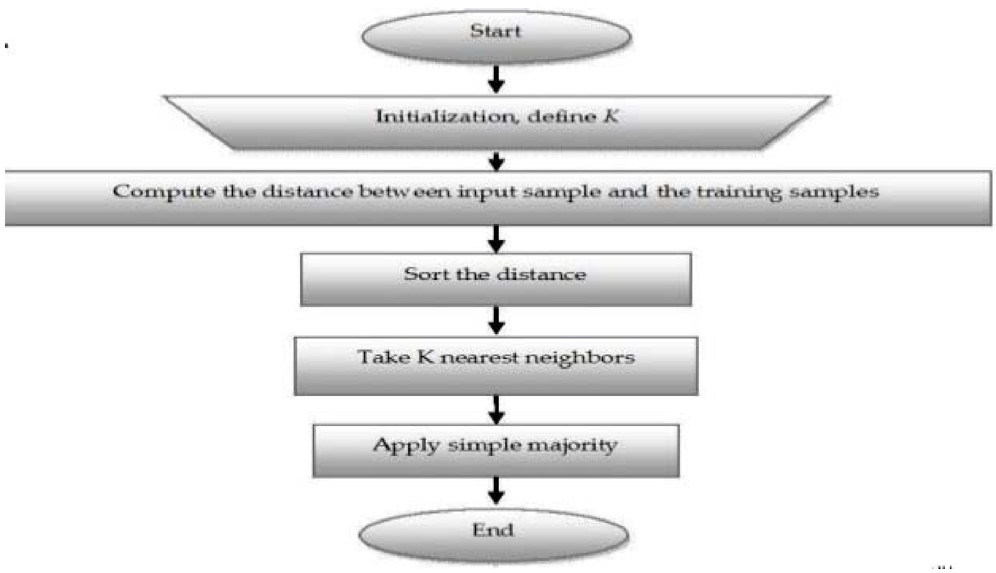


FIGURE 3. The Flowchart of KNN Method

Distances between input samples and training data can be calculated using a variety of methods, including Euclidean distance, Manhattan distance, and cosine similarity. The distance is calculated using the Euclidean method in this study. The Euclidean distance method, which has been modified to scale all attribute values to between 0 and 1, works well in domains where the attributes are all equally important to the outcome. Such domains, on the other hand, are the exception rather than the rule. Some attributes are irrelevant in most domains, while others are more important. The next step forward in instance-based learning is to incrementally learn the importance of each attribute by dynamically updating feature weights [14]. To calculate Euclidean distance, use

Formula (2), where W_{kj} is the value of TF-IDF from a term that has similarities to a term in the query, while W_{kq} is the TF-IDF value of a term in the query.

$$Sim(D,q) = \sqrt{\sum_{k=1}^{n} \left(W_{kq} - W_{kj}\right)^2} \tag{2}$$

Evaluation

In the validation system, this research employs the confusion matrix method. Confusion matrix is one of the important tools in the visualization method used in machine learning, which usually has two or more categories. Each matrix element displays the number of sample test data for the actual class, which is represented in rows, and the predicted class, which is represented in columns [15]. Table 1 illustrates examples of confusion matrix prediction results for two classes.

TABLE 1. Example of Confusion Matrix

		Actual Class	
		Class 1	**Class 2**
Predicted Class	**Class 1**	True Positive	False Negative
	Class 2	False Positive	True Negative

The correct prediction results are true positive (TP) and true negative (TN) values. The false positive value (FP) is the value predicted by the system to be class-1 but is actually class-2, whereas the false negative value (FN) is the value predicted by the system to be class-2 but is actually class-1. The accuracy value is obtained based on the correct prediction results divided by the total data or if based on Table 1, then the accuracy value is obtained by Formula (3).

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \tag{3}$$

RESULTS AND DISCUSSION

System Testing

This section went over the testing of the system that was built. The goal of this system is to detect tweets that have the potential to contain hate speech so that the user can determine whether the tweet is safe to post. Crawling, preprocessing, feature extraction, and training will be performed before the system can detect hate speech in tweets. As described in the previous section, the crawling process will be initiated to obtain the desired data. The crawled data are tweets with certain keywords that have been selected by an ITE law expert and several assistants. This selection process is carried out so that the tweet data obtained is in accordance with the hate speech cases that have occurred in Indonesia. The crawling process is carried out by utilizing the Twitter API as described in the previous section. The Twitter API has a maximum limit of 100 tweets for one crawling process. The total tweet data obtained from the crawling process is 973 tweets. After the tweet data was obtained, an ITE law expert and several research assistants selected the tweet data to discard tweets that were not relevant for this research. The total tweet data obtained from this selection process is 276 tweets, where this data will be further processed by the system. Several research assistants will manually label the collected data based on guidance from ITE law experts. For example, there is a tweet that says "*biksu penista Agama biksu itu botak semua itu mah biksu dukun*" so that tweet is labeled "*Pasal 156a KUHP*", because it is included in religious blasphemy. Table 2 shows the distribution of tweet data by class.

TABLE 2. Distribution of Tweet Data

Class	Number of Tweets
Pasal 156 KUHP	90
Pasal 156a KUHP	65
Pasal 45A ayat (2) jo. Pasal 28 ayat (2) UU ITE	68
Tidak Melanggar	53

After all the tweets have been labeled, the system will preprocess them. The first step in preprocessing is to remove symbols that frequently appear in tweet data, such as #, @, emoticons, and website links. The following step is case folding, which converts all the letters to lowercase. Table 3 shows examples of the result of symbol removal and case folding.

TABLE 3. Some Sample Tweets After Symbol Removal and Case Folding

No	Original Tweet	After Symbol Removal	After Case Folding
1	*@LsOwien Coba aja suruh tuh idi kacung who itu autopsi orang yang katanya mati karena coped ga akan berani*	*Coba aja suruh tuh idi kacung who itu autopsi orang yang katanya mati karena coped ga akan berani*	*coba aja suruh tuh idi kacung who itu autopsi orang yang katanya mati karena coped ga akan berani*
2	*Orang Kristen/Katolik mayoritas GAY.. RT @era_muslim: Paus : Gay Tidak Harus Dipinggirkan*	*Orang Kristen Katolik mayoritas GAY Paus Gay Tidak Harus Dipinggirkan*	*orang kristen katolik mayoritas gay paus gay tidak harus dipinggirkan*
3	*@VIVAcoid ohhh... jadi faham sampai sini... ternyata aktivis 98 biangkerok anti islam di negara ini.... yayaya*	*ohhh jadi faham sampai sini ternyata aktivis biangkerok anti islam di negara ini yayaya*	*ohhh jadi faham sampai sini ternyata aktivis biangkerok anti islam di negara ini yayaya*
4	*@iPatJnkyu Anjing lah mau jadi orang cina aja*	*Anjing lah mau jadi orang cina aja*	*anjing lah mau jadi orang cina aja*

The next step of preprocessing is stemming and stop word removal. The *Sastrawi 1.0.1* library was used in this study to perform stemming and stop words. Examples of stemming and stopword removal results are shown in Table 4. The tweets used as examples for stemming and stop word removal are the results of case folding in Table 3.

TABLE 4. Some Sample Tweets After Stemming and Stop word Removal

No	After Case Folding	After Stemming	After Stop word Removal
1	*coba aja suruh tuh idi kacung who itu autopsi orang yang katanya mati karena coped ga akan berani*	*coba aja suruh tuh idi kacung who itu autopsi orang yang kata mati karena coped ga akan berani*	*coba aja suruh tuh idi kacung who autopsi orang kata mati coped ga berani*
2	*orang kristen katolik mayoritas gay paus gay tidak harus dipinggirkan*	*orang kristen katolik mayoritas gay paus gay tidak harus pinggir*	*orang kristen katolik mayoritas gay paus gay pinggir*
3	*ohhh jadi faham sampai sini ternyata aktivis biangkerok anti islam di negara ini yayaya*	*ohhh jadi faham sampai sini nyata aktivis biangkerok anti islam di negara ini yayaya*	*ohhh jadi faham nyata aktivis biangkerok anti islam negara yayaya*
4	*anjing lah mau jadi orang cina aja*	*anjing lah mau jadi orang cina aja*	*anjing lah mau jadi orang cina aja*

The final step of preprocessing is tokenization. Tokenization will separate sentences into word by word based on spaces. Following the completion of the preprocessing, the tweet data will be subjected to feature extraction using the TF-IDF method before the training process begins. The system will automatically run the training after the feature extraction process is completed. The KNN method is used to carry out the training process. The *K* value used in the KNN method is 9. This value is derived from the results of the previous accuracy testing, which will be discussed in the following section.

After the training process is complete, the system is ready to detect new tweets entered by the user. Users are required to log in first with their respective Twitter accounts, before using this system. Figure 4 shows a view of the hate speech detection system built. Users can input new tweets, which will be detected through the text area provided by the system. The detection process will start after the user clicks the "*Check*" button, and the prediction results will be displayed below it. In Fig. 4, it can be seen that the prediction results are displayed in three columns. The first column contains the tweet that was entered, the second column contains the detection result, and the final column contains the maximum penalty that the user can receive if the tweet contains hate speech content. The system also provides a "*POST*" button, which is next to the detection result, where the button functions to post a tweet to our Twitter page.

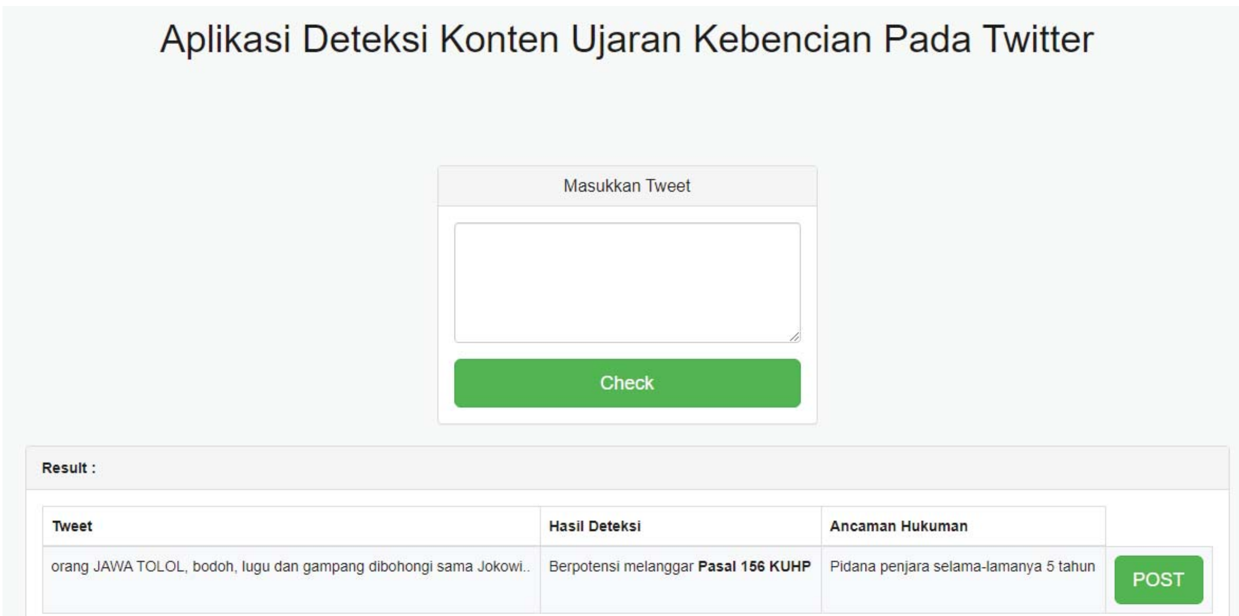


FIGURE 4. Hate Speech Detection System Display

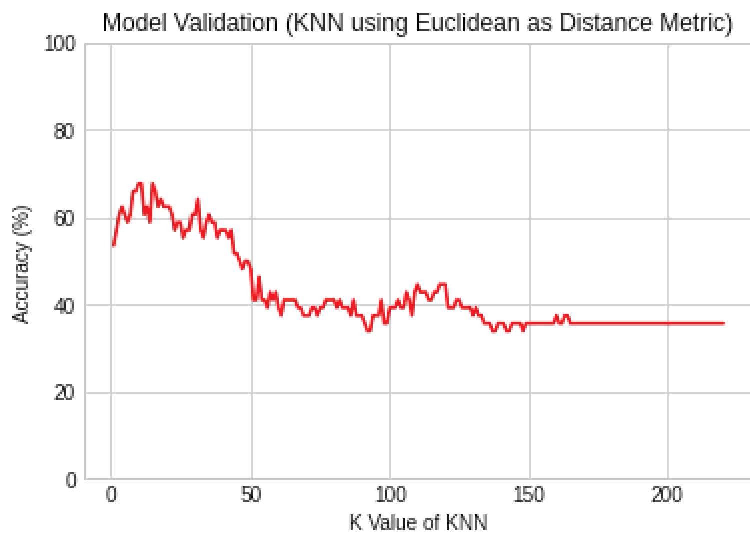
System Validation

The goal of system validation is to determine the level of accuracy of the system being built. Model validation and user validation are the two parts of system validation. The goal of model validation is to obtain the best accuracy value from the classifier used in this study, in this case the KNN method. The K value in the KNN method used in the training process will be determined by this validation. Model validation begins by dividing the owned tweet data into 80% training data and 20% testing data. Following that, the system will train using 80% of the training data, followed by a test using 20% of the testing data to determine the accuracy of the resulting classifier model. The ratio of the training and testing data was chosen because based on research conducted by Uçar et al [16], for a few datasets, the test rate that can be selected is between 10% and 20%. These processes are repeated from $K = 1$ to $K = 221$. The value of $K = 221$ was determined based on the total training data used, which was 80% of the 276 data. According to the experimental results of the model validation shown in Fig. 5, the highest accuracy is 67.86% with a K value of 10 for the KKN method using testing data of 20% of the total dataset owned. The confusion matrix from the model validation of the KNN method with a value of K=10 can be seen in Fig. 6. Based on the confusion matrix in Fig. 6, testing data that was classified correctly amounted to 38 tweets.

In addition to model validation for the KNN method that uses Euclidean as a distance metric, model validation is also carried out for other distance metrics, namely Manhattan and Cosine. It aims to compare and find out which distance metric has the best accuracy for this study. Based on the experimental results shown in Table 5, it can be seen that the validation model of the KNN method, with Manhattan as the distance metric, produces the highest accuracy of 57.14% with a value of K=6. Meanwhile, the highest accuracy for Cosine as a distance metric in the KNN method is 67.86% with a value of K=15.

TABLE 5. Comparison of Model Validation Results for Several Distance Metrics

No	Distance Metrics	Highest Accuracy	K value of KNN
1	Euclidean	67.86%	10
2	Manhattan	57.14%	6
3	Cosine	67.86%	15




```
Highest Accuracy= 67.86 %, with K= 10
```

FIGURE 5. The Experiment Results of The Model Validation

	Pasal 156 KUHP	Pasal 156a KUHP	Pasal 45A ayat (2) jo. Pasal 28 ayat (2) UU ITE	Tidak melanggar
Pasal 156 KUHP	17	2	0	1
Pasal 156a KUHP	2	8	1	0
Pasal 45A ayat (2) jo. Pasal 28 ayat (2) UU ITE	5	0	9	1
Tidak melanggar	3	1	2	4

FIGURE 6. The Confusion Matrix of The Model Validation

Based on the results of the model validation, the accuracy of the KKN method used in this study is quite good. However, the accuracy can still be improved by adding the number and variations of the tweet data used for the training process. Furthermore, based on the distribution of tweet data in Table 2, it is necessary to add tweet data to a certain class, so that the number of tweet data in each class is balanced. Another process that can be added to improve the accuracy of the classifier model is the process of converting slang words into standard words. Slang words are words that are not standardized but are commonly used in everyday life, especially activities on social media. Negative slang words were found in the tweet data used in this study, for example the words "*jancuk*", "*jancookk*", "*juancuuukk*", and "*jiannncuukkkk*". The four words actually have the same meaning, but are written differently, so the system will consider the four words to be different entities.

The second type of system validation is user validation. This validation is carried out by inputting 100 new tweets that are not in the dataset and are to be predicted by the system. The system's prediction results will be validated directly by ITE law experts. From the prediction results of 100 tweets that have been made, 77 of them are valid. Therefore, it can be concluded that the system accuracy based on user validation is 77%. The difference in accuracy between model validation and user validation lies in the amount of training data used. In model validation, the training process uses data from 80% of the total dataset owned. Meanwhile, in user validation, the training process uses 100% of the existing dataset. This proves that the more data used during the training process, the higher the accuracy produced.

CONCLUSION

Based on the model validation results, it is possible to conclude that the KNN method used in this research has the highest accuracy of 67.86%, with a K value of 10. Furthermore, based on the results of user validation, the hate speech content detection system developed has an accuracy of 77 percent, with 77 of the 100 tweets used for testing producing valid prediction results.

However, the system's accuracy can be improved in the future by increasing the number and variety of tweet data used in the training process. Furthermore, tweet data can be added in specific classes, ensuring that the number of tweets for each class is balanced. It is also necessary to include the process of converting slang words into standard words, so that duplication of words that have the same meaning can be eliminated.

ACKNOWLEDGMENTS

We would like to express our appreciation to LPPM Ubaya, who has supported this research, as well as other research assistants, in order for this research to be carried out successfully.

REFERENCES

1. B. Setiawan, "Indonesia Peringkat Lima Pengguna Twitter," Ministry of Communication and Informatics of Indonesia, 2 November 2012. [Online]. Available: https://kominfo.go.id/content/detail/2366/%20indonesia-peringkat-lima-penggunatwitter/0/sorotan_media. [Accessed 18 May 2021].
2. J. Mathias, "Hate Speech and Its Threat to Law Enforcement," The Indonesian Journal of International Clinical Legal Education, vol. 3, no. 1, 2021.
3. I. Saini, D. Singh dan A. Khosla, "QRS detection using K-Nearest Neighbor algorithm (KNN) and evaluation on standard ECG databases," Journal of Advanced Research, vol. 4, pp. 331-344, 2013.
4. I. Alfina, R. Mulia, M. I. Fanany dan Y. Ekanata, "Hate speech detection in the Indonesian language: A dataset and preliminary study," in 2017 International Conference on Advanced Computer Science and Information Systems (ICACSIS), Bali, 2017.
5. T. Febriana dan A. Budiarto, "Twitter Dataset for Hate Speech and Cyberbullying Detection in Indonesian Language," in 2019 International Conference on Information Management and Technology (ICIMTech), Jakarta, 2019.
6. J. Patihullah dan E. Winarko, "Hate Speech Detection for Indonesian Tweets Using Word Embedding And Gated Recurrent Unit," Indonesian Journal of Computing and Cybernetics Systems, vol. 13, no. 1, pp. 43-52, 2019.
7. D. A. N. Taradhita dan I. K. G. D. Putra, "Hate Speech Classification in Indonesian Language Tweets by Using Convolutional Neural Network," Journal of ICT Research and Applications, vol. 14, no. 3, pp. 225-239, 2021.

8. M. A. Kausar, V. S. Dhaka dan S. K. Singh, “Web crawler: a review,” International Journal of Computer Applications, vol. 63, no. 2, pp. 31-36, 2013.
9. M. R. Akbar, I. Slamet dan I. I. Handajani, “Sentiment analysis using tweets data from Twitter of Indonesian's Capital City changes using classification method support vector machine,” in International Conference On Science And Applied Science, Surakarta, 2020.
10. V. R. Prasetyo, “Searching Cheapest Product On Three Different E-Commerce Using K-Means Algorithm,” in International Seminar on Intelligent Technology and Its Application (ISITIA), Bali, 2018.
11. D. Wahyudi, T. Susyanto dan D. Nugroho, “Iimplementasi dan Analisis Algoritma Stemming Nazief & Adriani dan Porter Pada Dokumen Berbahasa Indoensia,” Jurnal Ilmiah SINUS, vol. 15, pp. 49-56, 2017.
12. V. R. Prasetyo, B. Hartanto dan A. A. Mulyono, “Penentuan Pembimbing Tugas Akhir Mahasiswa Jurusan Teknik Informatika Universitas Surabaya Dengan Metode Dice Coefficient,” TEKNIKA, vol. 8, no. 1, pp. 44-51, 2019.
13. V. Srividhya dan R. Anitha, “Evaluating preprocessing techniques in text categorization,” International Journal of Computer Science and Application Issue, pp. 49-51, 2010.
14. I. H. Witten, E. Frank, M. A. Hall dan C. J. Pal, Data Mining (Fourth Edition) Practical Machine Learning Tools and Techniques, Hamilton: University of Waikato, 2017.
15. V. R. Prasetyo dan E. Winarko, “Rating Of Indonesian Sinetron Based On Public Opinion In Twitter Using Cosine Similarity,” in International Conference On Science and Technology, Yogyakarta, 2017.
16. M. K. Uçar, M. Nour, H. Sindi dan K. Polat, “The Effect of Training and Testing Process on Machine Learning in Biomedical Datasets,” Hindawi Mathematical Problems in Engineering, vol. 2020, pp. 1-17, 2020.