ISBN: 978-1-7281-2379-0
IEEE Catalog Number: CFP19B09-USB

IEEE INDONESIA SECTION

ICSITech

PROCEEDING

2019 5th International Conference on Science in Information Technology (ICSITech)

October

23-24, 2019

Hotel Grand Inna Malioboro
Yogyakarta, Indonesia

Hosted by:

Jointly organized with:

2019 5th International Conference on Science in Information Technology (ICSITech)

October 23-24, 2019
Yogyakarta, Indonesia

COPYRIGHT AND REPRINT PERMISSION:

IEEE Catalog Number : CFP19B09-USB
ISBN : 978-1-7281-2379-0

Editor : Awang Hendrianto Pratomo, Andri Pranolo, Bagus Muhammad Akbar, Shoffan Saifullah, Ahmad Taufiq Akbar, Rochmat Husaini, Heriyanto, Andiko Putro Suryotomo, Vynska Amalia Permadi, Sylvert Prian Tahalea

Publisher : IEEE

Secretariat : Department of Informatics Engineering
Universitas Pembangunan Nasional Veteran Yogyakarta
Jln. Babarsari no. 2, Yogyakarta 55281
Special Region of Yogyakarta, Indonesia

PROCEEDING

2019 5th International Conference on Science in Information Technology (ICSITech)

"Embracing Industry 4.0 : Towards Innovation in Cyber Physical System"

Version: 2019-10-20

October 23-24, 2019
Yogyakarta, Indonesia

Introduction

Welcome to ICSITech 2019. Universitas Pembangunan Nasional "Veteran" Yogyakarta is honored to be the host of this year's International Conference on Science in Information Technology (ICSITech). The ICSITech is jointly organized with Universitas Ahmad Dahlan, Universitas Mulawarman, Universitas Pembangunan Nasional "Veteran" Yogyakarta, Universiti Teknologi Malaysia (UTM) Big Data Centre, Universitas Pendidikan Indonesia, Universitas Muhammadiyah Surakarta, Universiti Putra Malaysia, Universiti Malaysia Sabah, Universitas Budi Luhur, Politeknik Negeri Samarinda, Politeknik Negeri Padang, Universitas Negeri Malang, Universiti Teknikal Malaysia Melaka(UTeM), and ITSA University Colombia.

Since this is the Fifth conference, we wish to repeat the success of the four previous conferences. We do hope this annual conference will continue to be held in the next coming years (2020 in Malaysia, 2021 in Surakarta - Indonesia, etc.) with increasing quality. For this year's conference, we proudly present the theme of ICSITech 2019, "**Embracing Industry 4.0 : Towards Innovation in Cyber Physical System**". The theme is taken from our university identity as a university which consistently takes part in education and responds to the development of science, technology, art, society demands, and global change. It is also our university vision to be actively involved and acted as a leader and initiator in research and development in order to achieve academic excellence.

We are pleased to inform you that the ICSITech 2019 has been approved by IEEE for technical co-sponsorship; therefore, the papers which are accepted and presented will be further considered to be published in the IEEE Xplore Digital Library. I wish to extend a warm welcome to Prof. Dr. Wisnu Jatmiko, S.T., M.Kom., as IEEE Indonesia Section Chair. There are 118 papers from 17 countries submitted to the ICSITech 2019 with 48.28% acceptance ratio. Congratulations for all authors and presenters whose papers are accepted.

Thank you for choosing ICSITech 2019 and disseminating your research here.

Today, we are lucky to have two keynote speakers who will broaden our insights about Big Data Era in IT perspective. They will talk about their expertise and we do hope this event could bring many benefits, especially in the fields of education, industry, and society. We are honored for the presence of Prof. Dr. Leonel Hernandes (ITSA Colombia) and Assoc. Prof. Dr. Mohammad Shanuddin bin Zakaria (Universiti Kebangsaan Malaysia), thank you very much.

The previous conferences were held in Bandung-Indonesia and Melaka – Malaysia, respectively. This year, the ICSITech 2019 is taking place in Yogyakrta. Our city nickname is The City of Education. Yogyakarta is one of the favorite travel destinations, especially in Java, with many wonderful heritage tourism destinations and delightful culinary creations. Please enjoy your stay in Yogyakarta.

Finally, we would like to thank all of keynote speakers, participants, sponsors, associations, and partners for being a part of this conference. On behalf of the organizing committee, we wish to express our highest appreciation and sincere thanks to all of you who attend this event and we wish you have valuable discussion and networking. I also thank the committee for all efforts to make ICSITech 2019 successful.

Thank you.

General Chair
Dr. Awang Hendrianto Pratomo, S.T., M.T.
Department of Informatics Engineering
Universitas Pembangunan Nasional "Veteran" Yogyakarta, Indonesia

Welcome Message from Rector of Universitas Pembangunan Nasional Veteran Yogyakarta

Greetings to all distinguished guests, keynote speakers, and conference participants of the fifth (5th) International Conference on Science in Information Technology ICSITech 2019. It is our honor and pleasure to be the host of this year's ICSITech.

Ladies and gentlemen,
Universitas Pembangunan Nasional Veteran Yogyakarta has been taking part in organizing ICSITech since the first conference back in 2015. Since its inception ICSITech is jointly organized by Universitas Ahmad Dahlan, Universitas Mulawarman, Universitas Pembangunan Nasional "Veteran" Yogyakarta, Universiti Teknologi Malaysia (UTM) Big Data Centre, Universitas Pendidikan Indonesia, Universitas Muhammadiyah Surakarta, Universiti Putra Malaysia, Universiti Malaysia Sabah, Universitas Budi Luhur, Politeknik Negeri Samarinda, Politeknik Negeri Padang, Universitas Negeri Malang, Universiti Teknikal Malaysia Melaka(UTeM), and ITSA University Colombia. And this year, for the first time, our university got the opportunity to be the host of ICSITech 2019. In this occasion, I want to personally congratulate our Department of Informatics Engineering for their commitment and hard work to ensure the success of ICSITech 2019.

Ladies and gentlemen,
ICSITech 2019 was held to provide as an event for IT academic and IT expert to disseminate their knowledge on the development of computer science education and expand the network connection on the research activities. Furthermore we intend to make this conference as a motivation for researchers to publish their ideas about theory and application of IT for education, society, and industry in order to support the development and quality improvement of local, regional , and global researches, in line with Universitas Pembangunan Nasional Veteran Yogyakarta vision to be the pioneer of national development based on the spirit of patriotism.

Finally, on behalf of the organizing committee, we wish to express our highest appreciation and sincere thanks to all of our distinguished participants who attend this event and we wish you have valuable discussion and networking. Welcome to Yogyakarta and we hope all of our guests enjoying the services we provide.

Thank you.

Rector of Universitas Pembangunan Nasional Veteran Yogyakarta,
Dr. Mohamad Irhas Effendi, M.S.

Welcome Message from IEEE Indonesia Section

Prof. Dr.Eng Wisnu Jatmiko, SMIEEE
Chairman, IEEE Indonesia Section

Dr. Kurnianingsih, SMIEEE
Vice Chair, IEEE Indonesia Section

Dear Distinguished Guests, Colleagues, researchers, professionals, ladies and gentlemen, Good morning, a prosperous, warm, and spirited greeting.

On behalf of IEEE Indonesia Section, we would like to extend our warmest welcome to all keynote speakers, presenters, and participants to 2019 International Conference on Science in Information Technology (ICSITech). The conference theme is: "**Embracing Industry 4.0 : Towards Innovation in Cyber Physical System**".

ICSITech is an annual international conference technical co-sponsored by IEEE Indonesia Section and this year conference is hosted by Universiti Pembangunan Nasional "Veteran" Yogyakarta in collaboration with Universiti Teknologi Malaysia (UTM) Big Data Centre, Universitas Ahmad Dahlan, Universitas Mulawarman, Universitas Pendidikan Indonesi, Universitas Muhammadiyah Surakarta, Universiti Putra Malaysia, Universiti Malaysia Sabah, Universitas Budi Luhur, Politeknik Negeri Samarinda, Politeknik Negeri Padang, Universitas Negeri Malang, Universiti Teknikal Malaysia Melaka(UTeM), and ITSA University Columbia. The conference aims to bring together researchers and experts in information systems to share their ideas, experiences and insights.

IEEE Indonesia Section has conducted many activities over 32 years in Indonesia. In terms of collaboration, IEEE Indonesia section has a good and mutual relationship with ICT organizations, Industries, Government, Universities as well as the Community in Indonesia. IEEE Indonesia Section has contributed in about 58 different International conferences annually. As the fifth year of ICSITech, this conference shows its sustainability due to the hard work of the conference organizers, well organized conference and high-quality papers. We do hope in the near future some high-quality conferences will be continued and strengthened, so the result will give more benefit and positive impact to the human being, especially to Indonesian people.

In this occasion, I would also like to say welcome to Yogyakarta, which serves beautiful heritages, culture, with warm, polite and friendly people, a vibrant culture and lifestyle.
Finally, we do hope all of you will have enjoyable and valuable experience during this event. You may share your best knowledge in your area of research and professional activities.
Thank you.

Yogyakarta, 23rd October 2019
IEEE Indonesia Section

Organization

Steering Committee

- Munir (Universitas Pendidikan Indonesia)
- Dwi Hendratmo W. (Institut Teknologi Bandung)
- Satriyo Dharmanto (IEEE Indonesia Section)
- Shi-Jinn Horng (National Taiwan University Science and Technology, Taiwan)
- Hj. Siti Mariyam Shamsudin (Universiti Teknologi Malaysia)
- Amelia Ritahani Ismail (International Islamic University Malaysia)
- Rafal Drezewski (AGH University of Science and Technology)
- Rodziah Atan (Universiti Putra Malaysia)
- Rayner Alfred (Universiti Malaysia Sabah)
- Heui Seok Lim (Korea University)
- Goutam Chakraborty (Iwate Prefectural University)
- Didi Sukyadi (Universitas Pendidikan Indonesia)
- Tsukasa Hirashima (Hiroshima University)
- Halimah Badioze Zaman (Universiti Kebangsaan Malaysia)

Scientific Committee

- Agus Pratondo (Telkom University)
- Abderrafiaa Koukam (UTBM, France)
- Abdul Samad Shibghatullah (Universiti Teknikal Malaysia Melaka)
- Ade Gafar Abdullah (Universitas Pendidikan Indonesia)
- Adiwijaya (Telkom University)
- Ag Asri Ag Ibrahim (Universiti Malaysia Sabah)
- Agus Harjoko (Universitas Gadjah Mada)
- Agus Setiabudi (Universitas Pendidikan Indonesia)
- Agus Setiawan (Universitas Pendidikan Indonesia)
- Aji Wibawa (Universitas Negeri Malang)
- Alejandro Rosales (INAOE, Mexico)
- Amer Farea (Taiz University, Arab Saudi)
- Anca Ralescu (University of Cincinnati Ohio)
- Andria Arisal (Indonesia Institute of Science)
- Anhar Risnumawan (Politeknik Elektronika Negeri Surabaya)
- Anton Satria Prabuwono (King Abdulaziz University, Saudi Arabia)
- Anton Yudhana (Universitas Ahmad Dahlan)
- Arda Yunianta (King Abdulaziz University)
- Ari Barmawi (Telkom University)
- Armin Lawi (Hasanuddin University)
- Awang Hendrianto Pratomo (UPN "Veteran" Yogyakarta)
- Azhari SN (Universitas Gadjah Mada)
- Azizi Abdullah (Universiti Kebangsaan Malaysia)
- Azuraliza Abu Bakar (Universiti Kebangsaan Malaysia)
- Bana Handaga (Universitas Muhammadiyah Surakarta)
- Cepy Riana (Universitas Pendidikan Indonesia)
- Chin Kim On (Universiti Malaysia Sabah)
- Danial Hooshyar (Korea University)
- Dedi Rohendi (Universitas Pendidikan Indonesia)
- Deris Stiawan (Universitas Sriwijaya)
- Dewi Octaviani (Korea University)
- Didi Rosiyadi (Research Center for Informatics LIPI)
- Didin Wahyudin (Universitas Pendidikan Indonesia)
- Edi Kurniawan (Research Center for Informatics LIPI)
- Eko Aribowo (Universitas Ahmad Dahlan)

- Engkos Kosasih (Universitas Indonesia)
- Esa Prakasa (Research Center for Informatics LIPI)
- Faaizah Shahbodin (Universiti Teknikal Malaysia Melaka)
- Fajar Suryawan (Universitas Muhammadiyah Surakarta)
- Ford Lumban Gaol (Bina Nusantara University)
- Gede Pramudya Ananta (Universiti Teknikal Malaysia Melaka)
- Gunawan Ariyanto (Universitas Muhammadiyah Surakarta)
- Habib Shah (Universiti Tun Hussein Onn Malaysia)
- Halizah Basiron (Universiti Teknikal Malaysia Melaka)
- Hamzah bin Ahmad (Universiti Malaysia Pahang)
- Hanung Adi Nugroho (Universitas Gadjah Mada)
- Herlina Jayadianti (UPN Veteran Yogyakarta)
- Heru Supriyono (Universitas Muhammadiyah Surakarta)
- Hidayah Rahmalan (Universiti Teknikal Malaysia Melaka)
- Husni Thamrin (Universitas Muhammadiyah Surakarta)
- I Wayan Mustika (IEEE Inonesia Section)
- Ibrahim Ahmad (Universiti Teknikal Malaysia Melaka)
- Igi Ardiyanto (Universitas Gadjah Mada)
- Intan Ermahani A. Jalil (Universiti Teknikal Malaysia Melaka)
- Ito Wasito (Universitas Indonesia)
- Iwan Tri Riyadi Yanto (Universitas Ahmad Dahlan)
- Jazi Istiyanto (Universitas Gadjah Mada)
- Joey M. Suba (University of The Assumption)
- Kamarul Hawari bin Ghazali (Universiti Malaysia Pahang)
- Keng-Yap Ng (Universiti Putra Malaysia)
- Khabib Mustofa (Universitas Gadjah Mada)
- Lala Septem Riza (Universitas Pendidikan Indonesia)
- Lau Hui Keng (Universiti Malaysia Sabah)
- Leonel Hernandez (ITSA University)
- Lian Duan (New Jersey Institute of Technology, USA)
- Lili Wulandhari (Bina Nusantara University)
- Lizawati Salahuddin (Universiti Teknikal Malaysia Melaka)
- Maizatul Akmar Ismail (Universiti Malaya)
- Maman Abdurohman (Telkom University)
- Mar Yah Said (Universiti Putra Malaysia)
- Marzanah Abdul Jabar (Universiti Putra Malaysia)
- Masayu Leyla Khodra (Institut Teknologi Bandung)
- Masrah Azrifah Azmi Murad (Universiti Putra Malaysia)
- Mohammad Syafrullah (Universitas Budi Luhur, Jakarta)
- Mohd Hanafi Ahmad Hijazi (Universiti Malaysia Sabah)
- Mohd Sazali Khalid (Universiti Sains Malaysia)
- Mohd Shahizan Bin Othman (Universiti Teknologi Malaysia)
- Moslem Yousefi (UNITEN, Malaysia)
- Muh. Nursalman (Universitas Pendidikan Indonesia)
- Muhammad Hakim Othman (Universiti Teknikal Malaysia Melaka)
- Muhammad Said Hasibuan (IBI Darmajaya)
- Muhammad Zarlis (Universitas Sumatera Utara)
- Munir (Universitas Pendidikan Indonesia)
- Mustafa Kaiali (Mevlana University, Turkey)
- Nana Suryana Herman (Universiti Teknikal Malaysia Melaka)
- Nataniel Dengen (Universitas Mulawarman)
- Nazri Nawi (Universiti Tun Hussein Onn Malaysia)
- Ngo Hea Choon (Universiti Teknikal Malaysia Melaka)
- Nita Solehati (Chonbuk National University)
- Noel Lopes (Polytechnic of Guarda, Portugal)

- Noor Aida Husaini (Universiti Tun Hussein Onn Malaysia)
- Nor Fazlida Mohd Sani (Universiti Putra Malaysia)
- Nor Haslinda binti Ismail (Universiti Teknikal Malaysia Melaka)
- Noraini Che Pa (Universiti Putra Malaysia)
- Noraini Ibrahim (Universiti Tun Hussein Onn Malaysia)
- Norashikin Ahmad (Universiti Teknikal Malaysia Melaka)
- Norazah Mohd Nordin (Universiti Kebangsaan Malaysia)
- Norhalina Senan (Universiti Tun Hussein Onn Malaysia)
- Norhayati Harum (Universiti Teknikal Malaysia Melaka)
- Norhayati Mohd Ali (Universiti Putra Malaysia)
- Nurgiyatna (Universitas Muhammadiyah Surakarta)
- Nurul Amelina Nashruddin (Universiti Putra Malaysia)
- Nurul Azma Zakaria (Universiti Teknikal Malaysia Melaka)
- Oki Wicaksono (Universitas Gadjah Mada)
- Omar Al Jadaan (MHSU, Uni Arab Emirates)
- Omid Motlagh (CSIRO, Australia)
- Othman Mohd (Universiti Teknikal Malaysia Melaka)
- Ouri Wolfson (University of Illinois, USA)
- Parman Sukarno (Politeknik Pos Indonesia)
- Per Johan Runeson (System Lund University, Sweden)
- Prima Vitasari Hj. Purwadi (Universiti Malaysia Pahang)
- Raden Bagus Fajriya Hakim (Universitas Islam Indonesia)
- Rafah Mohamed Almuttairi (University of Babylon, Iraq)
- Rafal Drezewski (AGH UST, Poland)
- Ramadiani (Universitas Mulawarman)
- Rani Megasari (Universitas Pendidikan Indonesia)
- Rasim (Universitas Pendidikan Indonesia)
- Rayner Alfred (Universiti Malaysia Sabah)
- Reza Firsandaya Malik (Universitas Sriwijaya)
- Reza Pulungan (Universitas Gadjah Mada)
- Rifki Sadikin (Research Center for Informatics LIPI)
- Rimba Widhiana Ciptasari (Telkom University)
- Rinaldi Munir (Institut Teknologi Bandung)
- Riyanarto Sarno (Institut Teknologi Sepuluh Nopember)
- Rodina binti Ahmad (Universiti Malaya)
- Rodziah Atan (Universiti Putra Malaysia)
- Ronny Mardiyanto (Institut Teknologi Sepuluh Nopember)
- Rusli Abdullah (Universiti Putra Malaysia)
- Rusydi Umar (Universitas Ahmad Dahlan)
- Sa'adah Hassan (Universiti Putra Malaysia)
- Sabrina Ahmad (Universiti Teknikal Malaysia Melaka)
- Salfarina Abdullah (Universiti Putra Malaysia)
- Sarina Sulaiman (Universiti Teknologi Malaysia)
- Sarni Suhaila Rahim (Universiti Teknikal Malaysia Melaka)
- Shafaatunnur Hasan (Universiti Teknologi Malaysia)
- Shah Nazir (University of Swabi, Pakistan)
- Shahril bin Parumo (Universiti Teknikal Malaysia Melaka)
- Shaik Shakeel Ahamad (K.G. RCET, Hyderabad, India)
- Sharifah Sakinah Syed Ahmad (Universiti Teknikal Malaysia Melaka)
- Shi-Jinn Horng (NUST, Taiwan)
- Sigit Yuwono (Telkom University)
- Siti Mariyam Shamsuddin (Universiti Teknologi Malaysia)
- Siti Nurul Mahfuzah Mohamad (Universiti Teknikal Malaysia Melaka)
- Siti Sophiayati Yuhaniz (Universiti Teknologi Malaysia)
- Slamet Riyadi (Universitas Muhammadiyah Yogyakarta)

- Songhoua Xu (New Jersey Institute of Technology, USA)
- Sri Kusumadewi (Universitas Islam Indonesia)
- Suhirman (Universitas Teknologi Yogyakarta)
- Sukrisno Mardiyanto (Institut Teknologi Bandung)
- Sultan Noman Qasem (Taiz University, Arab Saudi)
- Sunardi (Universitas Ahmad Dahlan)
- Sunu Wibirama (Universitas Gadjah Mada)
- Sutarman (Universiti Malaysia Pahang)
- Suyanto (Telkom University)
- Suyoto (Universitas Atma Jaya Yogyakarta)
- Teguh Bharata Adji (Universitas Gadjah Mada)
- Teo Susnjak (Massey University, New Zealand)
- Tony Dwi Susanto (Institut Teknologi Sepuluh Nopember)
- Tutut Herawan (Universiti Malaysia Pahang)
- Ummi Rabaah Hashim (Universiti Teknikal Malaysia Melaka)
- Vimala Balakrishnan (Universiti Malaya)
- Waleed Ali Ahmed Abdullah (King Abdulaziz University, Arab Saudi)
- Wan Nurhayati Wan Ab Rahman (Universiti Putra Malaysia)
- Warusia Yassin (Universiti Teknikal Malaysia Melaka)
- Wawan Setiawan (Universitas Pendidikan Indonesia)
- Wendi Usino (Universitas Budi Luhur, Jakarta)
- Younes Saadi (Universiti Tun Hussein Onn Malaysia)
- Yudi Wibisono (Universitas Pendidikan Indonesia)
- Yusmadi Yah Jusoh (Universiti Putra Malaysia)
- Zakiah Ayop (Universiti Teknikal Malaysia Melaka)
- Zuwairie bin Ibrahim (Universiti Malaysia Pahang)

Organizing Committee

Patron
Rector of Universitas Pembangunan Nasional “Veteran” Yogyakarta

Advisor
- Vice Rector of Universitas Pembangunan Nasional “Veteran” Yogyakarta
- Dean of Faculty of Industrial Engineering Universitas Pembangunan Nasional “Veteran” Yogyakarta
- Head of Informatics Engineering Department Universitas Pembangunan Nasional “Veteran” Yogyakarta

General Chair
Awang Hendrianto Pratomo (Universitas Pembangunan Nasional “Veteran” Yogyakarta)

General Co-chair
- Haviluddin (Universitas Mulawarman)
- Aji Prasetya Wibawa (Universitas Negeri Malang)
- Bambang Yuwono (Universitas Pembangunan Nasional “Veteran” Yogyakarta)
- Hidayatullah Himawan, Universitas Pembangunan Nasional “Veteran” Yogyakarta

Secretary
- Oliver Samual Simanjuntak (Universitas Pembangunan Nasional “Veteran” Yogyakarta)
- Hari Prapcoyo (Universitas Pembangunan Nasional “Veteran” Yogyakarta)
- Rifki Indra Perwira (Universitas Pembangunan Nasional “Veteran” Yogyakarta)
- Sylvert Prian Tahalea (Universitas Pembangunan Nasional “Veteran” Yogyakarta)

Technical Program Chair
- Lala Septem Riza (Universitas Pendidikan Indonesia)
- Andri Pranolo (Universitas Ahmad Dahlan)
- Ramadiani (Universitas Mulawarman)
- Muhammad Syafrullah (Universitas Budi Luhur)
- Gunawan Ariyanto (Universitas Muhammadiyah Surakarta)
- Aji Prasetya Wibawa (Universitas Negeri Malang)
- Ummi Raba'ah Hashim (Universiti Teknikal Malaysia)
- Awang Hendrianto Pratomo (Universitas Pembangunan Nasional "Veteran" Yogyakarta)

Layout TPC
- Haviluddin Sukirno (Universitas Mulawarman)
- Yudi Wibisono (Universitas Pendidikan Indonesia)
- Hamdani (Universitas Mulawarman)
- Iwan Tri Riyadi Yanto (Universitas Ahmad Dahlan)
- Krisna Adiyarta (Universitas Budi Luhur)
- Adhi Prahara (Universitas Ahmad Dahlan)
- Oki Wicaksono (Universitas Mulawarman)
- Indra Riyanto (Universitas Budi Luhur)

Publication
- Simon Pulung Nugroho (Universitas Pembangunan Nasional "Veteran" Yogyakarta)
- Budi Santoso (Universitas Pembangunan Nasional "Veteran" Yogyakarta)
- Herlina Jayadianti (Universitas Pembangunan Nasional "Veteran" Yogyakarta)
- Shoffan Saifullah (Universitas Pembangunan Nasional "Veteran" Yogyakarta)

Treasury, Financial, & Sponsorship
- Wilis Kaswidjanti (Universitas Pembangunan Nasional "Veteran" Yogyakarta)
- Yuli Fauziah (Universitas Pembangunan Nasional "Veteran" Yogyakarta)
- Vynska Amalia Permadi (Universitas Pembangunan Nasional "Veteran" Yogyakarta)

Publicity & Website
- Dessyanto Boedi Prasetyo (Universitas Pembangunan Nasional "Veteran" Yogyakarta)
- Bagus Muhammad Akbar (Universitas Pembangunan Nasional "Veteran" Yogyakarta)
- Frans Richard Kodong (Universitas Pembangunan Nasional "Veteran" Yogyakarta)
- Ahmad Taufiq Akbar (Universitas Pembangunan Nasional "Veteran" Yogyakarta)

Registration, Certificate & Gift
- Herry Sofyan (Universitas Pembangunan Nasional "Veteran" Yogyakarta)
- Yuli Fauziah (Universitas Pembangunan Nasional "Veteran" Yogyakarta)
- Rochmat Husaini (Universitas Pembangunan Nasional "Veteran" Yogyakarta)

Technical, Facility & Logistic
- Agus Sasmito A. (Universitas Pembangunan Nasional "Veteran" Yogyakarta)\
- Andiko Putro Suryotomo (Universitas Pembangunan Nasional "Veteran" Yogyakarta)
- Riza Prapascatama Agusdin (Universitas Pembangunan Nasional "Veteran" Yogyakarta)

Event Management (Program & Protocol)
Herry Sofyan (Universitas Pembangunan Nasional "Veteran" Yogyakarta)

Keynote Speakers Biography

Leonel Eduardo Hernandez Collantes receive both his undergraduate degree in System Engineering and his postgraduate degree with Computer Networks specialization from Universidad del Norte, Barranquilla. He holds a masters degree in Strategic Management in Telecommunications from Universidad Internacional Iberoamericana. He has been a reviewer for scientific journal IJAIN since 2017 and an active member of IEEE Colombia section. He is an active researcher of the Telematics Engineering program, Faculty of Engineering, University Institution ITSA.

Assoc. Prof. Mohamad Shanudin Zakaria received his B.Sc. and M.Sc. titles from Northrop University, California. He holds a Ph.D. degree from University of Reading, United Kingdom. He is an Associate Professor in the Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia. He is part of a team taking a systemic approach and appllying ICT to the conservation of Lake Chini biodiversity and the improvement of its ecosystem. His research interest are application of System Thinking in ICT, architecting IT solutions, Business-IT alignment and ICT Governance.

Conference Schedule

Wednesday, October 23rd 2019

Time	Event
06:30 – 07:30	Hospitality & Registration Desk Open
07:30 – 09:00	Opening Ceremony
	1. National Anthem – Indonesia Raya
	2. Bela Negara Anthem
	3. Cultural Performance
	4. Welcome Address by ICSITech Chairperson
	5. Supporting Address by IEEE Indonesia Section
	6. Welcome Address by Rector of Universitas Pembangunan Nasional Veteran Yogyakarta
	7. Opening Ceremonial by Rector of Universitas Pembangunan Nasional Veteran Yogyakarta
	8. Photo Session
09:00 – 09:30	Coffee Break
09:30 – 12:00	Keynote Speech
	Keynote Speaker 1
	Keynote Speaker 2
	Moderator: Dr. Herlina Jayadianti, S.T., M.T.
12:00 – 13:00	Lunch and Prayer Time
13:00 – 15:00	Parallel Session 1
15:00 – 15:20	Coffee Break
15:20 – 17:20	Parallel Session 2
17:20 – 18.30	Break
18.30 – 21.00	Gala Dinner Session
	Best Paper Announcement
	Closing Ceremony
	Miscellaneous Information

Thursday, October 24th 2019

Yogyakarta Tour

Parallel Session Schedule

SAMAS Room – Artificial Intelligence Track

Time	Paper Title
13:00 – 13:20	**(#1570562661) Problem Event Extraction to Develop Causal Loop Representation from Texts** *Chaveevan Pechsiri, Intaka Piriyakul, Narongdech Keeratipranon*
13:20 – 13:40	**(#1570570061) An Analytical Study on Email Classification Using 10-Fold Cross-Validation** *Takorn Prexawanprasut, Piyanuch Chaipornkaew*
13:40 – 14:00	**(#1570572339) K-Nearest Neighbor (K-NN) based Missing Data Imputation** *Della Murbarani Prawidya Murti, Aji P Wibawa, Muhammad Akbar, Utomo Pujianto*
14:00 – 14:20	**(#1570570793) Seizure Detection Based on EEG Signals Using Katz Fractal and SVM Classifiers** *Inung Wijayanto, Achmad Rizal, Annisa Humairani*
14:20 – 14:40	**(#1570573510) Comparison of Performance Support Vector Machine Algorithm and Naïve Bayes for Diabetes Diagnosis** *Dominikus Boli Watomakin, Andi Wahju Rahardjo Emanuel*
14:40 – 15:00	**(#1570573730) Hybrid approach redefinition with SMOTE-CSELM in handling class imbalance problem** *Hartono Hartono, Erianto Ongko, Dahlan Adbullah*
15:00 – 15:20	**Coffee Break**
15:20 – 15:40	**(#1570574409) Adaptive Localization and Segmentation of Optic Disc using K-means and Active Contour** *Hanung Adi Nugroho, Augustine Herin Hutami, Eka Frannita, Rizki Nurfauzi*
15:40 – 16:00	**(#1570574423) A Scalable Approach for Improving Implementation of a Frequent Pattern Mining Algorithm using MapReduce Programming** *Md. Abed Hasan, Naima Hassan, Md. Hasibuzzaman, Mohammad Rezwanul Huq*
16:00 – 16:20	**(#1570583082) The Multi Layer Auto Encoder Neural Network (ML- AENN) for Encryption and Decryption of Text Message** *Achmad Gaffar, Arief Bramanto Wicaksono Putra, Rheo Malani*
16:20 – 16:40	**(#1570586194) An Effective Quantum Inspired Genetic Algorithm for Continuous Multiobjective Optimization** *Julisa Bana Abraham, Oyas Wahyunggoro, Noor Akhmad Setiawan*
16:40 – 17:00	**(#1570586348) Categorical Data Classification based on Fuzzy K-Nearest Neighbor Approach** *Heru C Rustamaji, Shalfa F Luhrie, Oliver Simanjuntak, Juwairiah Juwairiah, Bambang Yuwono*

PARANG TRITIS ROOM – Artificial Intelligence Track

Time	Paper Title
13:00 – 13:20	**(#1570568748) General Formula for Finding Optimize Hologram Reflector Dimension to Project Images on Screen** *Nur Ezha Vidawati, Syaiful Adala, Nugroho Pratama, Simon Nugroho*
13:20 – 13:40	**(#1570568879) Image-based processing for ripeness classification of oil palm fruit** *Anindita Septiarini, Hamdani Hamdani, Anita Ahmad Kasim*
13:40 – 14:00	**(#1570569209) Accurate and Low-cost Fingerprint Classification via Transfer Learning**

Time	Paper Title
	Latifah Listyalina, Ikhwan Mustiadi
14:00 – 14:20	**(#1570570191) Morphological Feature Extraction from Low Quality of Thick blood Smear Microphotographs** *Umi Salamah, Anto Nugroho, Ismail Eko Prayitno Rozi, Puji Budi Setia Asih*
14:20 – 14:40	**(#1570570214) Friendly and Progressive Visual Secret Sharing with Lossless Reconstruction** *Heri Prasetyo, Jing-Ming Guo, Hsia Chih-Hsien*
14:40 – 15:00	**(#1570570262) Blind Color Image Denoising Using Low Rank Matrix Minimization** *Heri Prasetyo, Umi Salamah, Wiranto Wiranto, Sari Sihwi, Winarno Winarno*
15:00 – 15:20	**Coffee Break**
15:20 – 15:40	**(#1570574888) Impact of Different Degree of Smoothing on Non-Local Means based Filter for Retinal Vessel Modeling** *Kartika Firdausy, Oyas Wahyunggoro, Hanung Adi Nugroho, Muhammad Bayu Sasongko, Risanuri Hidayat*
15:40 – 16:00	**(#1570584981) An Efficient Ear Recognition Based on Colour Texture and Edges Geometry** *Raad Ahmed Hadi*
16:00 – 16:20	**(#1570585570) Comparison of Egg Fertility Identification based on GLCM Feature Extraction using Backpropagation and K-means Clustering Algorithms** *Shoffan Saifullah, Vynska Amalia Permadi*
16:20 – 16:40	**(#1570585900) Plasmodium Parasites Detection in Digital Microscopic Images Containing of Luminance Noise** *Hanung Adi Nugroho, Rizki Nurfauzi, Eka Frannita*
16:40 – 17:00	**(#1570586492) Image Graph Matching Based on Region Adjacency Graph** *Akmal Akmal, Rinaldi Munir, Judhi Santoso*

PARANG KUSUMO ROOM – IoT and Big Data Track

Time	Paper Title
13:00 – 13:20	**(#1570562826) Implementation MQTT-SN Protocol on Smart City Application based Wireless Sensor Network** *M. Udin Harun Al Rasyid, Ferry Astika Saputra, Fadlul Fikri*
13:20 – 13:40	**(#1570586137) Smart Cane: Public Transportation Code Detection and Identification System for Visually Impaired** *Giva Andriana Mutiara, Gita Indah Hapsari, Kevin Suhandi*
13:40 – 14:00	**(#1570586147) Preventing Traffic Road Accident: Prototype of Monitoring Tires Failure Detection** *Giva Andriana Mutiara, Periyadi Periyadi, Mochamad Ghani Rachmatul Firdaus, Janfery Sinaga*
14:00 – 14:20	**(#1570586268) Feature Extraction of ECG Signals using Discrete Wavelet Transform and MFCC** *Siti Agrippina Alodia, Yusuf Risanuri Hidayat*
14:20 – 14:40	**(#1570586588) Anomaly-Based Intrusion Detection and Prevention Using Adaptive Boosting in Software-defined Network** *Rifki Indra Perwira, Yuli Fauziah, I Retya Mahendra, Dessyanto Boedi Prasetyo, Oliver Simanjuntak*
15:00 – 15:20	**Coffee Break**
15:20 – 15:40	**(#1570586677) Intrusion Detection System as Audit in IoT Infrastructure using Ensemble Learning and SMOTE Method**

Time	Paper Title
	Aldhi Ari Kurniawan, Heru Agus Santoso, Moch Arief Soeleman, Ahmad Zainul Fanani
15:40 – 16:00	**(#1570587190) Rule-Based Pronunciation Models to Handle OOV Words for Indonesian Automatic Speech Recognition System** *Fanda Yuliana Putri, Devin Hoesen, Dessi Lestari*
16:00 – 16:20	**(#1570567752) Performance of SVM in Classifying the Quartile of Computer Science Journals** *Sandika Maulana Putra, Aji P Wibawa, Triyanna Widiyaningtyas, Ilham A. E. Zaeni*
16:20 – 16:40	**(#1570570016) A Prediction Model for Human Happiness Using Machine Learning Techniques** *Piyanuch Chaipornkaew, Takorn Prexawanprasut*
16:40 – 17:00	**(#1570570293) Analysis Accuracy of Random Forest Model for Big Data - A Case Study of Claim Severity Prediction in Car Insurance** *Kartika Chandra Dewi, Hendri Murfi, Sarini Abdullah*
17:00 – 17:20	**(#1570573491) Performance Evaluation of Supervised Machine Learning Algorithms Using Different Data Set Sizes for Diabetes Prediction** *Melky Radja, Andi Wahju Rahardjo Emanuel*

BARON ROOM – Artificial Intelligence Track

Time	Paper Title
13:00 – 13:20	**(#1570586558) Weakening Feature Independence of Naïve Bayes Using Feature Weighting and Selection on Imbalanced Customer Review Data** *Reiza Adi Cahya, Fitra A. Bachtiar*
13:20 – 13:40	**(#1570586752) A Semiautomatic Sperm Cell Data Annotator for Convolutional Neural Network** *Priyanto Hidayatullah, Tati Erawati Rajab, Rinaldi Munir, Anggraini Barlian*
13:40 – 14:00	**(#1570586813) Improving Imbalanced Dataset Classification Using Oversampling and Gradient Boosting** *Nur Heri Cahyana, Siti Khomsah, Agus Sasmito Aribowo*
14:00 – 14:20	**(#1570574779) Comparison Performance of C4.5, Naïve Bayes and K- Nearest Neighbor in Determination Drug Rehabilitation** *Islamiyah Islamiyah, Nataniel Dengen, Anisa Afiyah, Medi Taruk*
14:20 – 14:40	**(#1570586627) Eyeball Movement Detection Using Sector Line Distance Approach and Learning Vector Quantization** *Gusti Pangestu, Fitra A. Bachtiar*
14:40 – 15:00	**(#1570586955) Enumeration of Plasmodium Parasites on Thin Blood Smear Digital Microscopic Images** *Alifia Prananda, Hanung Adi Nugroho, Igi Ardiyanto*
15:00 – 15:20	**Coffee Break**
15:20 – 15:40	**(#1570587166) Arnold Transformed Position Power First Mapping (AT- PPFM) for Secret Digital Image** *Wilis Kaswidjanti, Hidayatulah Himawan, Afra Oryza Mursita Dewi, Mangaras Yanu Florestiyanto, Rifki Indra Perwira*
15:40 – 16:00	**(#1570586958) Wound Classifications Of Breast Tissues with Electrical Impedance Spectroscopy (EIS): Comparison of LVQ and GA-LVQ** *Yoke K Arbawa, R. Adhitya Pisefty, Fitra A. Bachtiar*
16:00 – 16:20	**(#1570570315) Vowel Sound Analysis in the Indonesian Language using Multilevel Wavelet Entropy** *Risanuri Hidayat, Achmad Rizal, Agus Bejo, Sujoko Sumaryono*

Time	Paper Title
16:20 – 16:40	**(#1570570348) Analysis Accuracy of XGBoost Model for Multiclass Classification - A Case Study of Applicant Level Risk Prediction for Life Insurance** *Widya Fajar Mustika*
16:40 – 17:00	**(#1570583031) Cosine Similarity for Title and Abstract of Economic Journal Classification** *Putri Yuni Ristanti, Aji P Wibawa, Utomo Pujianto*
17:00 – 17:20	**(#1570586600) Handling Imbalance Issue in Hate Speech Classification using Sampling-based Methods** *Rathpisey Heng, Teguh Bharata Adji*

Table of Contents

Problem Event Extraction to Develop Causal Loop Representation from Texts 1
Chaveevan Pechsiri; Intaka Piriyakul; Narongdech Keeratipranon

Implementation MQTT-SN Protocol on Smart City Application based Wireless Sensor Network 7
M. Udin Harun Al Rasyid; Ferry Astika Saputra; Fadlul Fikri

Performance of SVM in Classifying the Quartile of Computer Science Journals 13
Sandika Maulana Putra; Aji P Wibawa; Triyanna Widiyaningtyas; Ilham A. E. Zaeni

General Formula for Finding Optimize Hologram Reflector Dimension to Project Images on Screen 18
Nur Ezha Vidawati; Syaiful Adala; Nugroho Pratama; Simon Pulung Nugroho

Image-Based Processing for Ripeness Classification of Oil Palm Fruit 23
Anindita Septiarini; Hamdani Hamdani; Anita Ahmad Kasim

Accurate and Low-cost Fingerprint Classification via Transfer Learning 27
Latifah Listyalina; Ikhwan Mustiadi

A Prediction Model for Human Happiness Using Machine Learning Techniques 33
Piyanuch Chaipornkaew; Takorn Prexawanprasut

An Analytical Study on Email Classification Using 10-Fold Cross-Validation 38
Takorn Prexawanprasut; Piyanuch Chaipornkaew

Morphological Feature Extraction from Low Quality of Thick blood Smear Microphotographs 44
Umi Salamah; Anto Nugroho; Ismail Eko Prayitno Rozi; Puji Budi Setia Asih

Friendly and Progressive Visual Secret Sharing with Lossless Reconstruction 49
Heri Prasetyo; Jing-Ming Guo; Hsia Chih-Hsien

Blind Color Image Denoising Using Low Rank Matrix Minimization 55
Heri Prasetyo; Umi Salamah; Wiranto Wiranto; Sari Sihwi; Winarno Winarno

Analysis Accuracy of Random Forest Model for Big Data - A Case Study of Claim Severity Prediction in Car Insurance 60
Kartika Chandra Dewi; Hendri Murfi; Sarini Abdullah

Vowel Sound Analysis in the Indonesian Language using Multilevel Wavelet Entropy 66
Risanuri Hidayat; Achmad Rizal; Agus Bejo; Sujoko Sumaryono

Analysis Accuracy of XGBoost Model for Multiclass Classification - A Case Study of Applicant Level Risk Prediction for Life Insurance 71
Widya Fajar Mustika, Hendri Murfi, Yekti Widyaningsih

Seizure Detection Based on EEG Signals Using Katz Fractal and SVM Classifiers 78

Inung Wijayanto; Achmad Rizal; Annisa Humairani

K-Nearest Neighbor (K-NN) based Missing Data Imputation 83
Della Murbarani Prawidya Murti; Aji P Wibawa; Muhammad Akbar; Utomo Pujianto

Comparison of Performance Support Vector Machine Algorithm and Naïve Bayes for Diabetes Diagnosis 89
Dominikus Boli Watomakin; Andi Wahju Rahardjo Emanuel

Hybrid approach redefinition with SMOTE-CSELM in handling class imbalance problem 95
Hartono Hartono; Erianto Ongko; Dahlan Adbullah

Adaptive Localization and Segmentation of Optic Disc using K-means and Active Contour 100
Hanung Adi Nugroho; Augustine Herin Hutami; Eka Frannita; Rizki Nurfauzi

A Scalable Approach for Improving Implementation of a Frequent Pattern Mining Algorithm using MapReduce Programming 106
Md. Abed Hasan; Naima Hassan; Md. Hasibuzzaman; Mohammad Rezwanul Huq

Comparison Performance of C4.5, Naïve Bayes and K-Nearest Neighbor in Determination Drug Rehabilitation 112
Islamiyah Islamiyah; Nataniel Dengen; Anisa Afiyah; Medi Taruk

Impact of Different Degree of Smoothing on Non-Local Means based Filter for Retinal Vessel Modeling 118
Kartika Firdausy; Oyas Wahyunggoro; Hanung Adi Nugroho; Muhammad Bayu Sasongko; Risanuri Hidayat

Cosine Similarity for Title and Abstract of Economic Journal Classification 123
Putri Yuni Ristanti; Aji P Wibawa; Utomo Pujianto

The Multi Layer Auto Encoder Neural Network (ML-AENN) for Encryption and Decryption of Text Message 128
Achmad Gaffar; Arief Putra, ABWP; Rheo Malani

An Efficient Ear Recognition Based on Colour Texture and Edges Geometry 134
Raad Ahmed Hadi, Loay Edwar George

Comparison of Egg Fertility Identification based on GLCM Feature Extraction using Backpropagation and K-means Clustering Algorithms 140
Shoffan Saifullah; Vynska Amalia Permadi

Plasmodium Candidate Detection on Thin Blood Smear Images with Luminance Noise Reduction 146
Hanung Adi Nugroho; Rizki Nurfauzi; Eka Frannita

Smart Cane: Public Transportation Code Detection and Identification System for Visually Impaired 150
Giva Andriana Mutiara; Gita Indah Hapsari; Kevin Suhandi

Preventing Traffic Road Accident: Prototype of Monitoring Tires Failure Detection 156
Giva Andriana Mutiara; Periyadi Periyadi; Mochamad Ghani Rachmatul Firdaus; Janfery Sinaga

An Effective Quantum Inspired Genetic Algorithm for Continuous Multiobjective Optimization 161
Julisa Bana Abraham; Oyas Wahyunggoro; Noor Akhmad Setiawan

Feature Extraction of ECG Signals using Discrete Wavelet Transform and MFCC 167
Siti Agrippina Alodia Yusuf; Risanuri Hidayat

Categorical Data Classification based on Fuzzy K-Nearest Neighbor Approach 171
Heru C Rustamaji; Shalfa F Luhrie; Oliver Simanjuntak; Juwairiah Juwairiah; Bambang Yuwono

Image Graph Matching Based on Region Adjacency Graph 176
Akmal Akmal; Rinaldi Munir; Judhi Santoso

Weakening Feature Independence of Naïve Bayes Using Feature Weighting and Selection on Imbalanced Customer Review Data 182
Reiza Adi Cahya; Fitra A. Bachtiar

Anomaly-Based Intrusion Detection and Prevention Using Adaptive Boosting in Software-defined Network 188
Rifki Indra Perwira; Yuli Fauziah; I Retya Mahendra; Dessyanto Boedi Prasetyo; Oliver Simanjuntak

Handling Imbalance Issue in Hate Speech Classification using Sampling-based Methods 193
Rathpisey Heng; Teguh Bharata Adji

Eyeball Movement Detection Using Sector Line Distance Approach and Learning Vector Quantization 199
Gusti Pangestu; Fitra A. Bachtiar

Intrusion Detection System as Audit in IoT Infrastructure using Ensemble Learning and SMOTE Method 205
Aldhi Ari Kurniawan; Heru Agus Santoso; Moch Arief Soeleman; Ahmad Zainul Fanani

A Semiautomatic Sperm Cell Data Annotator for Convolutional Neural Network 211
Priyanto Hidayatullah; Tati Erawati Rajab; Rinaldi Munir; Anggraini Barlian

Improving Imbalanced Dataset Classification Using Oversampling and Gradient Boosting 217
Nur Heri Cahyana; Siti Khomsah; Agus Sasmito Aribowo

Enumeration of Plasmodium Parasites on Thin Blood Smear Digital Microscopic Images 223

Alifia Prananda; Hanung Adi Nugroho; Igi Ardiyanto

Wound Classifications Of Breast Tissues with Electrical Impedance Spectroscopy (EIS): Comparison of LVQ and GA-LVQ 229
Yoke K Arbawa; R. Adhitya Pisefty; Fitra A. Bachtiar

Comparative Performance Analysis Between Software-Defined Networks and Conventional IP Networks 235
Leonel Hernandez; Genett Jimenez; Andri Pranolo; Carlos Uc Rios

Arnold Transformed Position Power First Mapping (AT-PPFM) for Secret Digital Image 241
Wilis Kaswidjanti; Hidayatulah Himawan; Afra Oryza Mursita Dewi; Mangaras Yanu Florestiyanto; Rifki Indra Perwira

Rule-Based Pronunciation Models to Handle OOV Words for Indonesian Automatic Speech Recognition System 246
Fanda Yuliana Putri; Devin Hoesen; Dessi Lestari

Performance Evaluation of Supervised Machine Learning Algorithms Using Different Data Set Sizes for Diabetes Prediction 252
Melky Radja; Andi Wahju Rahardjo Emanuel

Comparison of Performance Support Vector Machine Algorithm and Naïve Bayes for Diabetes Diagnosis

Dominikus Boli Watomakin, Andi Wahju Rahardjo Emanuel
Magister Teknik Informatika, Universitas Atma Jaya Yogyakarta, Indonesia 55281
j1mmywatomakin@gmail.com, andi.emanuel@uajy.ac.id

Abstract - **Handling in the health sector has now developed a lot in terms of information technology. Many studies in the field of information technology that help in accelerating the performance of management of a health agency or from a work of health workers who require fast and good decision making. In this study a comparison of algorithms was used to diagnose diabetes, which had been used from many previous studies. Support vector machines and naïve bayes become comparison algorithms carried out in this study. The purpose of this study was to look at the performance of the two algorithms and help health workers in better decision making. The level of accuracy, precision, sensivity and specificity of the two algorithms will be the main focus of this research. Comparisons were made using a diabetes dataset taken from the National Institute of Diabetes and Digestive and Kidney Diseases with a total sample data of 768 sample data. From the results of calculations and comparisons of support vector machine algorithms have a better average value compared to the naïve Bayes algorithm.**

Keyword – Diabetes, Machine Learning, Support Vector Machine, Naïve Bayes, WEKA

I. INTRODUCTION

Diabetes is a serious chronic disease caused by the pancreas does not produce enough insulin, or when the body does not effectively use the insulin it produces. Diabetes is one of the many diseases suffered by the people of Indonesia and from the data obtained from the Data and Information Center of the Indonesian Ministry of Health in 2018, the number of diabetics in Indonesia is in 4th position, in 2000 diabetics numbered 8.4 million and it is estimated that by 2030 it will increase to 21.3 million if the lifestyle adopted by the community has not been properly managed [1]. The situation faced in Indonesia by looking at the results of the 1995-2001 Household Health Survey (SKRT) and Riskesdas 2007 showed that non-communicable diseases, one of which was diabetes mellitus, was the leading cause of death in Indonesia. In 2007, 59.5% of the causes of death in Indonesia were non-communicable diseases. In addition, the percentage of deaths from non-communicable diseases also increased from year to year, namely 41.7% in 1995, 49.9% in 2001, and 59.5% in 2007 [1].

From the problems that exist, and seeing health as a very important thing in people's lives, the medical side needs the right treatment to be able to diagnose diabetes quickly and accurately and also ongoing education and patient self-management support is very important to prevent acute complications and reduce the risk of long-term complications [2]. This makes the role of health workers have challenges in making good, fast and appropriate decisions to diagnose diabetes that patients suffer [2]. Often decision making affects external factors and causes mistakes that can obscure existing decisions [3]. To minimize errors in decision making by health professionals, the hospital or other health access utilizes information technology which is currently developing very rapidly [4]. Machine learning techniques are one of the techniques used in previous research to diagnose diseases [5]. Machine learning is a sub-field of artificial intelligence that solves real-world problems by providing learning skills to computers without additional programming [6]. Machine learning algorithms used in the medical world can be very beneficial in decision making [5]. Besides being able to help health workers, this method can increase public confidence in using this health service.

In connection with the narration described above, this study will compare two algorithms that have often been used in the health sector for decision making by health workers. The algorithm that will be compared is the support vector machine and naïve Bayes. The dataset used is the diabetes dataset taken from the National Institute of Diabetes and Digestive and Kidney Diseases. All patients in the dataset were women who had at least 21 years of age from Pima Indian descent. While the number of sample data there are as many as 768 sample data. Comparisons are made using the WEKA version 3.8 tool which is a tool for developing machine learning and data mining techniques [7]. By comparing these two algorithms you can find out the level of accuracy, precision, sensitivity and specificity so that it can be a recommendation and help health workers in making decisions in diagnosing diabetes.

II. LITERATURE REVIEW

In handling health workers, information technology is available in the health sector and has been highly developed. Many studies have produced important information technology contributions to help doctors, nurses or other medical officers make decisions that are often inappropriate [8]. Machine learning is one of the methods of information technology that is applied in the health sector. Many types of research on machine learning in the health sector have been developed. For example in research [9], reviewing prediction models (machine learning) in medicine and research [10], using the Bayesian hidden Markov Model method with a Gaussian Mixture Clustering approach to model changes in DNA copy numbers throughout the genome.

There are also several other studies that review or review various methods that exist in machine learning in terms of application in the health sector in the treatment of diabetes. Research [6], reviewing Machine Learning techniques in the classification of Diabetes and Cardiovascular Diseases (CVD) using Artificial Neural Networks (ANNs) and Bayesian Networks (BNs), while in research [11], doing a systematic review to predict diabetes using machine learning and data mining, in this study SVM became a very reliable algorithm. Research [12], use adjudication to measure errors in the assessment of Diabetic Retinopathy based on individual graders and majority decisions and to train an improved

automated algorithm for DR grading. To determine the accuracy of predictions of diabetes, research [13], assess the association between the Hypertriglyceridemic Waist (HW) phenotype and type 2 diabetes in adults in Korea, while the study [14], also increases the level of accuracy by developing extreme learning machines (ELM) to recognize the presence of hypoglycemia.

Handling of diabetes through alternative ways also helped develop the potential with the diagnosis of type 2 diabetes using soft computing methods [15]. Then research [16], they sought to streamline Machine learning algorithms for predicting effective outbreaks of chronic diseases in communities that often suffer from the disease. From previous research, many machine learning techniques have been carried out. Algorithms that exist in machine learning are widely applied to assist handling in the health field. What is done in this study is more than comparing algorithms that are often used in existing research. The author takes the SVM and naïve bayes algorithm for comparison using the diabetes dataset. In addition to doing a comparison of the two algorithms, this is done by the author to recommend one method that can be used by medical parties to treat diabetes, where as has been explained in the introduction above that one of the biggest mortality rates is caused by this disease.

III. MATERIALS AND METHOD

In this study, to compare the two algorithms, there are materials and methods used and explained in this chapter.

A. Datasets

The diabetes dataset used comes from the National Institute of Diabetes and Digestive and Kidney Diseases. This dataset is a dataset used to diagnose whether patients have diabetes, based on certain diagnostic measurements in the dataset. Some constraints are placed on selecting a larger database instance. Sample data contained in this dataset were 768 sample data of 21-year-old female sex from Pima Indian descent.

B. Method

To get good results in the comparisons made, this research went through several systematic steps. These stages are:

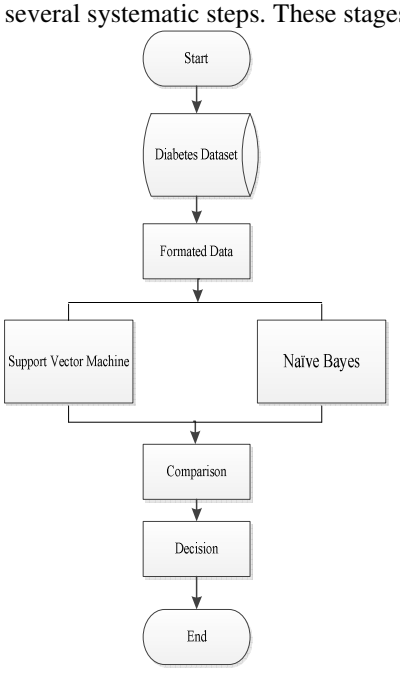

Figure 1. Research Methodology

The flow of this research begins with finding and preparing a diabetes dataset with format.csv, then with the dataset formatting will be carried out according to the tool that will be used to reduce errors while the process is running and in this study, the tool used is WEKA version 3.8 tool.

The next step is testing with two algorithms, namely support vector machine and naïve Bayes. After testing the next stage, a comparative analysis of the two algorithms will be carried out by looking at the level of accuracy with criteria consisting of precision, accuracy, specificity, and sensitivity. From the results of the comparative analysis carried out the results can be obtained which can be concluded which algorithm is better than the two algorithms in terms of accuracy for case studies identifying diabetes using machine learning.

C. Classification Algorithms

In this study, there are two algorithms that are carried out the comparative analysis. Briefly, in this subchapter, the two algorithms are explained:

1) Support Vector Machine

Support vector machine (SVM) is a supervised machine learning model that is known for its outstanding performance in pattern recognition and classification tasks with high dimensional data. SVM is a learning system that utilizes hypothesis space in the form of linear functions in a feature space that has high dimensions and is trained with learning algorithms based on optimization theory by implementing learning bias derived from statistical learning theory [17].

2) Naïve Bayes

Bayesian classification is a process that estimates the probability of new observations belonging to a predetermined category, using a probability model that is defined according to Bayes' theory [18]. In the naïve Bayes algorithm has similarities in general as follows:

$$P(C_j|X) = \frac{P(X|C_j)}{P(X)} . P(C_j) \quad (1)$$

Where $P(C_j|X)$ is the probability of an unknown observation X belongs to the category C_j and called posteriori probability, $P(X|C_j)$ is probability, given a category C_j, unknown observations belonging to this category, $P(C_j)$ is the previous probability of observation X unknown to be observed in categories C_j, and $P(X)$ is the previous probability of an unknown observation X is the same for each category C_j [18]. Bayes's theorem is often also developed considering the validity of the law of total probability, as follows:

$$P(C_j|X) = \frac{P(X|C_j)}{\sum_{i=1}^{n} P(C_i|X)} . P(C_j) \quad (2)$$

In the naïve bayes theory, the classification process requires a number of clues to be able to determine what class is suitable for the sample being analyzed. Therefore, the Bayes theorem above can be adjusted as follows:

$$P(C|F_1, \dots, F_n) = \frac{P(C)P(F_{1,}, \dots, F_n|C)}{P(F_{1,}, \dots, F_n)} \quad (3)$$

From equation (3), variable C is expressed as a class, then variable $F_1, \dots, F_n$ is a characteristic of the instructions needed to make a classification. Then the formula can be explained as

an opportunity for the entry of samples with certain characteristics in class *C* (posterior) which is the opportunity for the emergence of class *C* multiplied by the chance of the emergence of sample characteristics in class *C* (likelihood), divided by opportunities for the emergence of global sample characteristics (evidence). Therefore, formula (4) can be simply written as follows:

$$posterior = \frac{prior\ x\ likelihood}{evidence} \quad (4)$$

Evidence values are always fixed for each class in one sample. The value of the posterior will be compared with the posterior value of the other classes to determine to what class a sample will be classified.

D. Experiment

This stage is the stage of comparison analysis of two algorithms, namely support vector machine and naïve bayes by using WEKA version 3.8 tool. Experiments were carried out using datasets originating from the National Institute of Diabetes and Digestive and Kidney Diseases. The dataset used is a dataset of diabetes which uses sample data of 768 sample data. To see the performance and significance of the two algorithms, the experimental calculation will divide the dataset into four parts in the WEKA tool using a split which is 20%, 40%, 60% and 80% respectively. Then, the results of the four sections will be seen in four aspects of measurement, namely Accuracy, Precision, Sensitivity and Specificity.

Accuracy: is the level of similarity of predictions produced by the system with the actual value.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (5)$$

Precision: is the level of accuracy produced by the system with the actual value of information.

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

Sensitivity: is the level of identification produced by the system with the actual value.

$$Sensitivy = \frac{TP}{TP + FN} \quad (7)$$

Specificity: is the proportion produced by the system with the actual value.

$$Specificity = \frac{TN}{TN + FP} \quad (8)$$

Where True Positive *(TP)* is the positive proportion in data sets classified as positive, True Negative *(TN)*, the negative proportion in the data set classified as negative, False Positive *(FP)* is a negative proportion in the data set that is classified as positive, False Negative *(FN)* represents the negative proportion in the data set classified negative.

From the four aspects above, an accuracy calculation method will be used in the concept of data mining, namely using confusion matrix. Confusion matrix provides decisions obtained in training and testing, the confusion matrix provides an assessment of the performance of classification based on the object correctly or incorrectly. Confusion matrix contains actual and predicted information on the classification system. Confusion matrix can be seen in table 1.

Table 1. Confusion matrix

Confusion matrix		**True value**	
		True	**False**
Prediction Value	**True**	TP (True Positive) Correct result	FP (False Positive) Unexpected result
	False	FN (False Negative) Missing result	TN (True Negative) Correct absence of result

In table 1 above, the values of *TP* (true positive) and *TN* (true negative) indicate the level of classification accuracy. Generally, the higher the value of *TP* and *TN*, the better the classification level of accuracy, precision, and recall. If the output prediction label is true and the true value is false (false) it is called a false positive (*FP*). Whereas if the output label prediction is false (false) and the true value is true (true), this is called false negative (*FN*).

IV. RESULTS AND DISCUSSION

This section will explain the results of calculations and comparisons of SVM and naïve bayes algorithms that have been done. As explained in the previous chapter the dataset used is the diabetes dataset originating from the National Institute of Diabetes and Digestive and Kidney Diseases with a total sample data of 768 sample data.

The comparison of the two algorithms uses the WEKA version 3.8 tool, from which the confusion matrix will be obtained from four data splits for the two algorithms. The results obtained can be seen in table 2 and table 3.

Table 2. Confusion Matrix from SVM

Confusion matrix		***True Values***	
		True	***False***
Split 20%	True	104	107
	False	41	362
Split 40%	True	85	70
	False	34	272
Split 60%	True	57	48
	False	18	184
Split 80%	True	27	22
	False	9	96

Table 3. Confusion Matrix from naïve Bayes

Confusion matrix		***True Values***	
		True	***False***
Split 20%	True	122	89
	False	73	330
Split 40%	True	94	61
	False	56	250
Split 60%	True	62	43
	False	31	171
Split 80%	True	30	19
	False	16	89

From the two results of the confusion matrix above, it will be taken to do the next step, which is calculating and seeing the comparison of SVM and naïve bayes algorithms using formula accuracy (5), precission (6), sensivity (7) and specificity (8). The results of the calculation of each performance and comparison of two algorithms can be seen in table 4 and table 5 below.

Table 4. Comparison of Performance of SVM Algorithms

Split Data	Accuracy	Precision	Sensivity	Specificity
Split 20%	76%	49%	72%	77%
Split 40%	77%	55%	71%	80%
Split 60%	79%	71%	76%	79%
Split 80%	80%	80%	75%	81%

From table 4 above, it can be seen that when the dataset is divided into WEKA tools with four split percentages, it can be seen that the four measurement aspects tend to increase with the percentage of data split 20% to the percentage of data split 80%. On the measurement aspects of accuracy, the increase occurred from the percentage of data split 20% by 76% and increased to the percentage of data split 80% by 80%. However, in the other three aspects of measurement, there are still ups and downs in the calculation results. More details can be seen in the graph in figure 2 below.

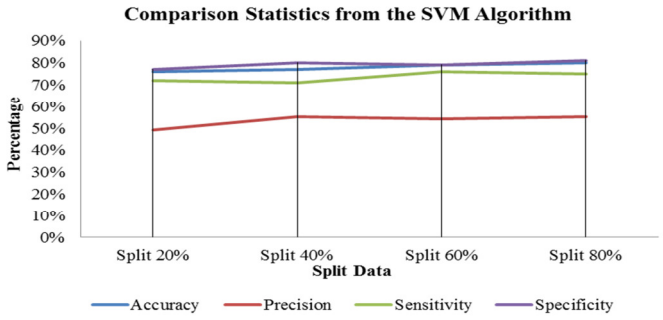

Figure 2. Comparison Statistics from the SVM Algorithm

In Figure 2, shows comparison statistics of the SVM algorithm. In the visualization that is displayed, it can be seen that the increase in the four aspects tends to increase but not too significant.

Table 5. Comparison of the Performance of the Naïve Bayes Algorithm

Split Data	Accuracy	Precision	Sensivity	Specificity
Split 20%	74%	58%	63%	79%
Split 40%	75%	61%	63%	80%
Split 60%	76%	59%	67%	80%
Split 80%	77%	61%	65%	82%

From table 5 above, it can be seen that the performance comparison on the naïve bayes algorithm, when the dataset is divided into WEKA tools with four split percentages, it can be seen that the four measurement aspects still show an increasing tendency. More details can be seen in the graph in figure 3 below.

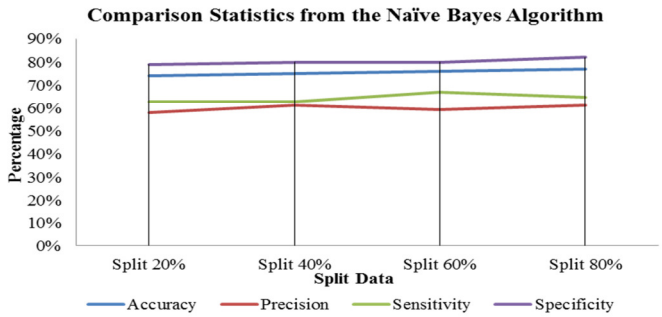

Figure 3. Comparison Statistics from the Naïve Bayes Algorithm

In Figure 3, shows comparative statistics from the naïve Bayes algorithm. In the visualization displayed, it can be seen that the increase in the four aspects tends to increase slightly but not too significantly. In accuracy measurement aspects and specificity measurement aspects, it can be seen an increase from the percentage of data split 20% to the percentage of data split 80%, but in the precision measurement aspect there is an increase in the percentage of data split 20% to the percentage of data split 40% and then the percentage of data split 60 % and in the percentage of data split 80% looks up. The measurement aspect of sensivity looks the same in the percentage of data split 20% and the percentage of data split 40% which then increases in the percentage of data split 60% and looks down on the percentage of data split 80%.

V. CONCLUSION

From the results of calculations and comparisons of SVM and naïve bayes algorithms using the diabetes dataset from the National Institute of Diabetes and Digestive and Kidney Diseases with a total sample data of 768 sample data carried out four times the percentage of data split beforehand that the two algorithms have advantages and differences in each aspect. Both algorithms have different statistical comparison averages and can be seen in the following figure 4.

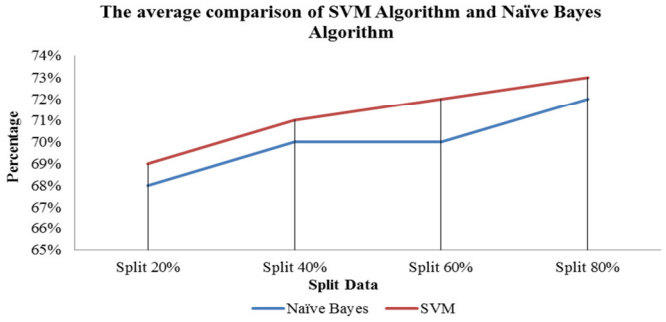

Figure 4. The average comparison of SVM Algorithm and Naïve Bayes Algorithm

The average comparison of the two algorithms shown in Figure 4 above shows that the SVM algorithm looks superior compared to the naïve bayes algorithm and tends to increase in every percentage split done. In the SVM algorithm the average percentage of split 20% by 69% compared to naïve bayes algorithm of 68%. Then in the percentage of 40% split SVM algorithm is 71% compared to naïve bayes algorithm of 70%. At 60% split percentage SVM algorithm is 72% compared to naïve bayes algorithm of 70%, and in the percentage of 80% split SVM algorithm is 73% compared to naïve bayes algorithm of 72%.

This research cannot be said to be perfect, therefore it is proposed for further research to be able to use deep learning techniques and make comparisons with other algorithms so that in terms of helping diagnosing diabetes can be even better..

REFERENCES

[1] K. N. K. R. Maula, "Hari Diabetes Sedunia Tahun 2018," 2018.

[2] D. Care and S. S. Suppl, "Introduction: Standards of medical care in Diabetes 2018," *Diabetes Care*, vol. 41, no. Supplement 1, pp. S1–S2, 2018.

[3] A. Bari, R. A. Khan, and A. W. Rathore, "Medical errors; causes, consequences, emotional response and resulting behavioral change," *Pakistan J. Med. Sci.*, vol. 32, no. 3, pp. 523–528, 2016.

[4] Y. K. Alotaibi and F. Federico, "The impact of health information technology on patient safety," *Saudi Med. J.*, vol. 38, no. 12, pp. 1173–1180, 2017.

[5] R. C. Deo, "Machine Learning in Medicine," *Circulation*, vol. 132, no. 20, pp. 1920–1930, 2015.

[6] B. Alić, L. Gurbeta, and A. Badnjević, "Machine learning techniques for classification of diabetes and cardiovascular diseases," *2017 6th Mediterr. Conf. Embed. Comput. MECO 2017 - Incl. ECYPS 2017, Proc.*, no. June, 2017.

[7] S. Sakr *et al.*, "Comparison of machine learning techniques to predict all-cause mortality using fitness data: The Henry Ford exercIse testing (FIT) project," *BMC Med. Inform. Decis. Mak.*, vol. 17, no. 1, pp. 1–15, 2017.

[8] T. Zheng and Y. Zhang, "A Machine Learning-based Framework to Identify Type 2 Diabetes through Electronic Health Records," *Commun. Comput. Inf. Sci.*, vol. 731, pp. 451–458, 2017.

[9] H. O. Alanazi, A. H. Abdullah, and K. N. Qureshi, "A Critical Review for Developing Accurate and Dynamic Predictive Models Using Machine Learning Methods in Medicine and Health Care," *J. Med. Syst.*, vol. 41, no. 4, 2017.

[10] G. Manogaran, V. Vijayakumar, R. Varatharajan, P. Malarvizhi Kumar, R. Sundarasekar, and C. H. Hsu, "Machine Learning Based Big Data Processing Framework for Cancer Diagnosis Using Hidden Markov Model and GM Clustering," *Wirel. Pers. Commun.*, vol. 102, no. 3, pp. 2099–2116, 2018.

[11] I. Kavakiotis, O. Tsave, A. Salifoglou, N. Maglaveras, I. Vlahavas, and I. Chouvarda, "Machine Learning and Data Mining Methods in Diabetes Research," *Comput. Struct. Biotechnol. J.*, vol. 15, pp. 104–116, 2017.

[12] J. Krause *et al.*, "Grader Variability and the Importance of Reference Standards for Evaluating Machine Learning Models for Diabetic Retinopathy," *Ophthalmology*, vol. 125, no. 8, pp. 1264–1272, 2018.

[13] G. Luo, "Automatically explaining machine learning prediction results: A demonstration on type 2 diabetes risk prediction," *Heal. Inf. Sci. Syst.*, vol. 4, no. 1, pp. 1–9, 2016.

[14] P. Samant and R. Agarwal, "Machine learning techniques for medical diagnosis of diabetes using iris images," *Comput. Methods Programs Biomed.*, vol. 157, pp. 121–128, 2018.

[15] B. J. Lee and J. Y. Kim, "Identification of type 2 diabetes risk factors using phenotypes consisting of anthropometry and triglycerides based on Machine Learning," *IEEE J. Biomed. Heal. Informatics*, vol. 20, no. 1, pp. 39–46, 2016.

[16] M. Chen, Y. Hao, K. Hwang, L. Wang, and L. Wang, "Disease Prediction by Machine Learning over Big Data from Healthcare Communities," *IEEE Access*, vol. 5, pp. 8869–8879, 2017.

[17] E. A. Shams and A. Rizaner, "A novel support vector machine based intrusion detection system for mobile ad hoc networks," *Wirel. Networks*, vol. 24, no. 5, pp. 1821–1829, 2018.

[18] P. Tsangaratos and I. Ilia, "Comparison of a logistic regression and Naïve Bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size," *Catena*, vol. 145, pp. 164–179, 2016.

Certificate of Appreciation

This is certify that

Andi Wahju Rahardjo Emanuel

has participated as

Presenter

In 2019 5th International Conference
on Science in Information Technology (ICSITech)

Yogyakarta - Indonesia, 23-24 October 2019

General Chair of ICSITech 2019

Dr. Awang Hendrianto Pratomo, S.T., M.T.

Jointly organized with:

Telkom University · UTM · UPM · UTeM · ITSA