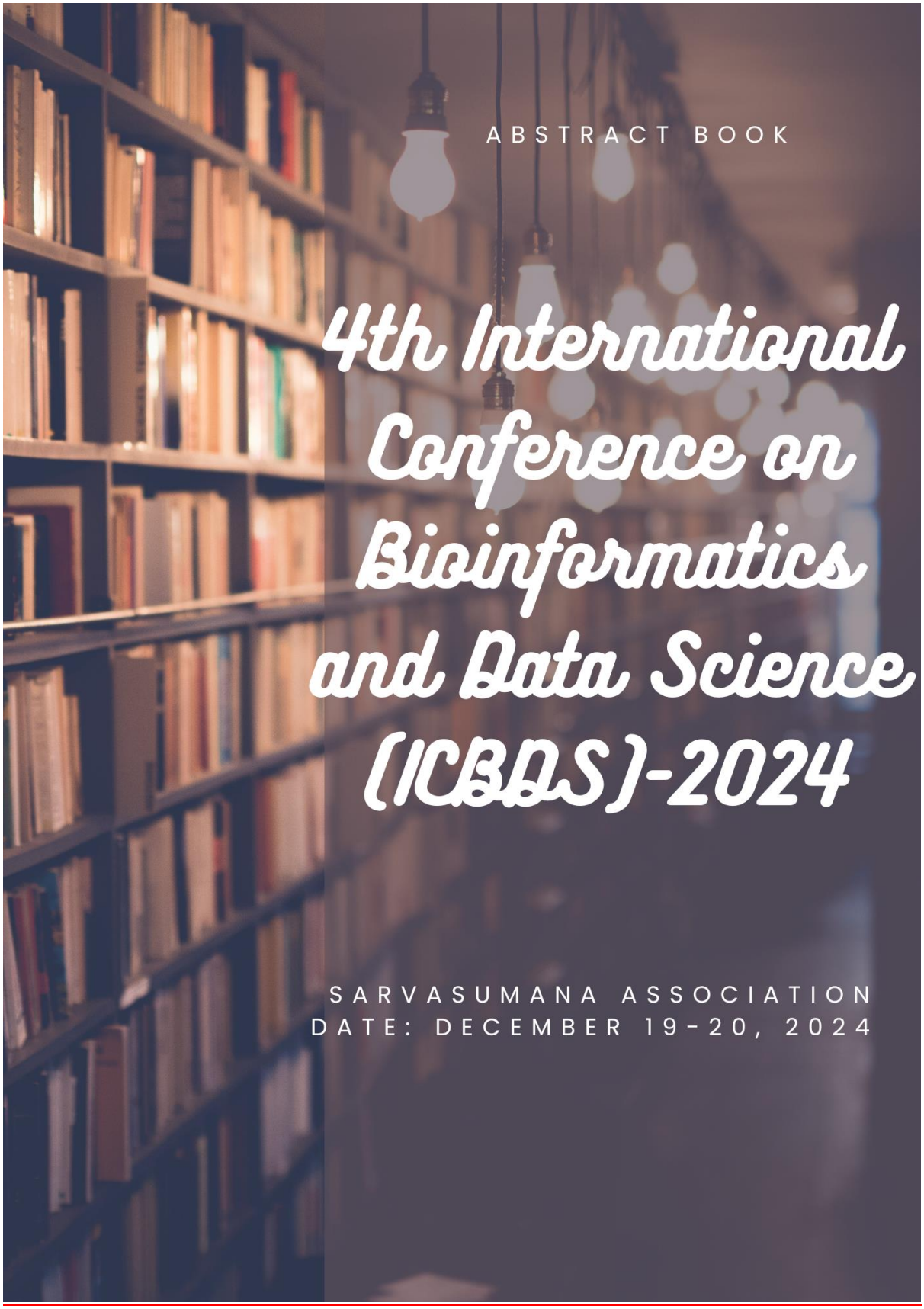

ABSTRACT BOOK

4th International Conference on Bioinformatics and Data Science (ICBDS)-2024

SARVASUMANA ASSOCIATION

DATE: DECEMBER 19-20, 2024

International Journal of Public Mental Health and Neurosciences
ISSN: 2394-4668 (Special Issue)

ORGANIZER	Sarvasumana Association, Bengaluru, India

ORGANIZING COMMITTEE

PATRON	Mrs. Padmashree Murthy
Chairman	Dr. R. Somashekhar
Secretary	Dr. Kirthi S Jawalkar
Convener	Dr. Preenon Bagchi
Organizing Joint Secretary and Co-Convener	Dr.Rajesh TS Mr. Shylesh Murthy IA
Student Volunteer Members	Ms. Jimema Paul Ms. Milan Khandelwal
National Advisory Committee	Dr. G. Dhanalakshmi Dr. Satish Sankaye Dr. Prashanth N Suravajhala Dr. Ayyakkannu Selvaraj
International Advisory Committee	Dr. Kathryn Lane Rossi, USA Dr. Richard Hill, Australia Dr. KongFatt Wong-Lin, Ireland

Sarvasumanâ Association

Reg. No.: JNRS79-2014-15
Registered Scientific and Industrial Research Organization (SIRO)
By Department of Scientific and Industrial Research (DSIR),
Govt. of India.

सर्वसुमना

#168 A, 1st cross,
1st Block, 3rd Phase,
BSK 3rd stage,
Bangalore-560085

WELCOME

On behalf of the Sarvasumana Association, it is my privilege to welcome you to the 4th International Conference on Bioinformatics and Data Science 2024. This conference serves as a vibrant platform for researchers, scholars, and professionals from around the globe to connect, collaborate, and share insights on the ever-evolving domains of bioinformatics and data science.

As we stand on the frontier of transformative discoveries, the integration of data-driven technologies with biological sciences continues to open new avenues for innovation. This year's conference will feature leading experts, insightful presentations, and groundbreaking research that aim to bridge the gap between science and technology, fostering solutions to global challenges in healthcare, biotechnology, and beyond.

We are honored to have esteemed speakers and participants joining us, and I encourage you to make the most of this opportunity to learn, exchange ideas, and build lasting collaborations.

I wish you a productive and enriching conference and look forward to the knowledge and connections that will emerge over the course of this event.

Thank you, and welcome once again!

With best wishes.

Padmashree murthy

Padmashree Murthy.
President
Sarvasumana Association

SARVASUMANA ASSOCIATION BANGALORE

+91-9448685484
https://sarvasumana.in

Vasishth Genomics Research Lab (P) Ltd®
Recognized as Start-Up by DIPP and Start-Up Karnataka
CIN: U73100KA2021PTC150940

FROM CHAIRMAN'S DESK

Respected Guests, Distinguished Speakers, Esteemed Researchers, and Participants, It is with immense pride and joy that I welcome you all to the **4th International Conference on Bioinformatics and Data Science 2024**. As the Chairman of this conference, it is both an honor and a privilege to witness this gathering of brilliant minds, passionate about advancing knowledge in the fields of **bioinformatics** and **data science**.

We stand at an exciting juncture where interdisciplinary approaches are driving scientific progress. Bioinformatics, coupled with data science, continues to revolutionize areas like genomics, healthcare, drug discovery, and personalized medicine. This conference serves as a vital platform for **sharing innovative ideas, presenting pioneering research, and fostering collaborations** that address global scientific and societal challenges.

I express my deepest gratitude to our keynote speakers, panelists, and participants who have joined us from across the globe. Your contributions are the heart of this event, and I am confident that your discussions, presentations, and interactions will inspire new perspectives and groundbreaking discoveries.

I would also like to extend my appreciation to the organizing committee, sponsors, and volunteers for their tireless efforts in making this event a reality.

Together, let us embark on this exciting journey of knowledge exchange, exploration, and discovery. I wish you all a very productive and inspiring conference.

Thank you, and let us begin this event with enthusiasm and collaboration!

Yours sincerely,

Dr. R. Somashekhar,
Director (Research)

Sl.No	Plenary Talks
1	Natural Language Processing: Challenges & Opportunities
	Dr. Satish Sankaye
2	Bridging the gap between Biologists and data scientists: Where are we heading? My reflections from our case studies
	Dr. Prashanth N Suravajhala
3	Stroke Risk Assessment using LIME to Interpret Deep Learning Predictions
	Dr. Ayyakkannu Selvaraj
4	Dynamical systems perspective of perceptual decision-making and metacognition
	Dr. Kongfatt Wong-Lin
5	Computational Approaches in Prodrug Discovery: A Bioinformatics Perspective
	Dr. G. Dhanalakshmi

Sl.No	Abstracts
1	Fruit Peel Biofertilizer: Harnessing Spoilage for Sustainable Agriculture
	Omkar Baliram Agale
2	Comparative Antibacterial Analysis of Natural Products Against Staphylococcus aureus Isolated from Fresh Wound Samples
	Akash Ajay Choudhari
3	Optimizing Testosterone through Nutrition: A Comprehensive Review
	Ronit Rohidas Savale
4	Exploring the Intersection of Psychology, Parapsychology, and Philosophy
	K.S. SRI RAJESHWARI DEVI
5	Identification and characterization of the Tau protein mutation and their role in protein structure and function
	Sreekanya Roy
6	Circular RNA: An Important Biomolecule to Control Gastric Cancer
	JIT MONDAL
7	Safed Musli (*Chlorophytum borivilianum* L.) tuber Mediated Green synthesis of Silver Nanoparticles and their characterization
	Dr. Rupali Taur
8	Developing Resilient Speech Emotion Recognition Systems through Deep Learning and Audio Augmentation for Enhanced Emotion Detection
	Irfan Abdullatif Chaugule
9	Machine Learning Driven Subtype and Mutation Classification in Gliomas Using Transcriptomic Data
	Shreeya Pahune
10	Unravelling Plasmodium falciparum's Life Cycle: Insights from Integrated Transcriptomic and Machine Learning Approaches
	Shreeya Pahune
11	Prediction of Parkinson's Disease Using Multimodal Machine Learning and Integration Techniques
	Shreeya Pahune
12	Network-Driven Drug Repurposing: Insights from Graph Neural Networks and Recommendation Models
	Shreeya Pahune
13	Towards Effective MCQ Generation: Semantic and Machine Learning Approaches"
	Jeevan Pralhad Tonde
14	Enhancing Data Security with DNA Methylation-Based Encryption Techniques
	Rajat Pandey
15	Big data in bioinformatics
	Sourabh Maid
16	Interactions between long non-coding RNAs and proteins to elucidate the mechanism of neurodegenerative disease onset
	Aniruddha Biswas
17	Revisiting Sushruta Samhita and Investigating the Potential Link between its Treatment Modalities by Predicting Study using Bioinformatics
	Jain Vaibhavi Ravindra

18	Usage of Natural Language Processing in extracting the invaluable insights From Sushruta Samhita
	Jain Vaibhavi Ravindra
19	Exploring Spiritual Wisdom: NLP-Based Knowledge Extraction from the Works of Sri Sri Thakur Anukulchandra
	Arpita Kharat
20	In Silico Study Of *Syzygium Cumini* Components With Hmg-Coa Reductase For Potential Anti-Obesity Effects
	Navya Raj A R

PLENARY TALKS

NATURAL LANGUAGE PROCESSING: CHALLENGES & OPPORTUNITIES

Dr. Satish R. Sankaye

Associate Professor, Dr.G.Y.Pathrikar College of Computer Science & Information Technology, MGM University, Chhatrapati Sambhaji Nagar, Maharashtra.431005

Email id: ssankaye@mgmu.ac.in

Abstract:

Language has been humanity's most powerful tool for communicating meaningful information effectively. In the 21st century, technological advancements in voice, text, and image processing have further enhanced this capability. Natural Language Processing (NLP) stands as a transformative technology that enables machines and devices to understand and process human language efficiently. By combining computational linguistics with statistical models, machine learning, and deep learning techniques, NLP has revolutionized human-computer interaction. The global NLP market is poised for significant growth, with an expected revenue of US$ 439.86 billion by 2030, growing at a rate of 40.4%. The Asia-Pacific region is projected to lead this growth, with India playing a dominant role in the expansion.

NLP offers astounding benefits across multiple domains. It enables the analysis of large datasets, increases documentation efficiency and accuracy, and provides objective and data-driven insights. Additionally, NLP enhances customer experiences by analyzing social media conversations, open-ended surveys, and online reviews, helping businesses better understand user sentiment and behavior. However, despite these advantages, NLP faces significant challenges. These include language differences, word-sense disambiguation, precision issues, tone of voice and inflection interpretation, and ambiguity in human communication. Addressing these challenges is critical to unlocking the full potential of NLP and ensuring its continued growth and impact.

BRIDGING THE GAP BETWEEN BIOLOGISTS AND DATA SCIENTISTS: WHERE ARE WE HEADING? MY REFLECTIONS FROM OUR CASE STUDIES

Dr. Prashanth N Suravajhala

Founder, Bioclues.org

Principal Scientist/Associate Professor, Amrita University, Kerala

E mail: prash@Bioclues.org

Post COVID-19 times has ushered a fierce competition to deliver, be it vaccine or funding or publication. As researchers, we have a fair conception to be guided by reasons not emotions amid 'publish or perish' adage. On the other hand, multitasking research and publishing has become a noticeable goal, but combining these tasks over time has become the need of the hour. In today's reserved funding situation, many early/mid-career researchers face a daunting task to establish and develop their research programs, for example starting their own labs crowdsourcing or obtaining funds from their previous associations/host institutions and publishing it. But to what extent are we trying to preserve the fairness or integrity of science? I would like to draw your attention to 'Hippocratic Oath for Scientists', which would ensure the research vitality in the best interests of science to sustain excellence. Towards this, the talk would delve on how the three Cs, viz. Consistency, Continuity and Credibility augur well for a successful open organization. This would invariably bring successful Collaborations, Convergence, and importantly Control over mind to the fore. The growth of an individual or organization depends on fostering commitment to open culture, net neutrality and universal access to information in education and science fields. So, it is the Collaborative index (C-index) that matters. Are we ready?

The second part of my talk would be on the CA Prostate Consortium of India (CAPCI) that I founded. As Prostate Cancer (PCa) burgeones, our recent comprehensive bioinformatics analysis confirmed some characteristic known or unknown mutations from a WES study native to India. Although our study shows characteristic mutations in certain genes, an

assay comprising multiple biomarkers that are differentially expressed could be attempted in the future. If this is successful, the number of biomarkers developed will depend on their validation in a large cohort of patients and the translation of these findings to clinical practice. From our pilot study, we believe, it is decisive to understand the inherent genes and mutations responsible for PCa in India. Furthermore, an attempt was made to develop a conceptual framework for research particularly in propagating information on the causal genes and mutations responsible for PCa. Although the work was limited to a small number of samples studied, we deem this pilot work would have an impending role in understanding mutations that are of particular interest to Indian genealogy. We ask whether a small compendium of sequential data we have could infer pathogenic mutations. We will discuss our impending exome and transcriptome case studies in making

Taken Together, Are we bridging the gap between the stakeholders?

STROKE RISK ASSESSMENT USING LIME TO INTERPRET DEEP LEARNING PREDICTIONS

Dr. Ayyakkannu Selvaraj

University Department of information and communication Technology (UDICT), MGM University, Chh.Sambhajinagar-431003

Stroke Risk Assessment using LIME to Interpret Deep Learning Predictions Ayyakkannu Selvaraj1, Subastri Arumugam2, Sharvari Tamane1, Rethinavalli3 1University Department of Information and communication Technology, MGM University 2 Department of Computer Science and business system Saranathan Engineering College, Tiruchirappalli, Tamil Nadu, India 3 Department of MCA, Shrimati Indira Gandhi College, Tiruchirappalli, India Abstract Stroke is a significant global health concern, influenced by a complex interplay of demographic, clinical, lifestyle, and socioeconomic risk factors. With approximately 13 million cases and 5.5 million fatalities reported annually by the World Stroke Organization, understanding and mitigating these risks is crucial for effective prevention. Hypertension remains the most significant modifiable risk factor, while dietary choices and physical activity also play critical roles in influencing stroke risk. This study examines the intricate relationship between various health-related factors and stroke risk using a dataset obtained from Kaggle, consisting of 3,254 participants aged 18 years and older. The dataset includes ten features encompassing demographic information, health conditions, and lifestyle choices, with stroke serving as the target variable. Significant associations were observed between smoking status and increased stroke rates, particularly among former smokers. Additionally, the analysis indicated a pronounced incidence of stroke among individuals aged 60 to 80, highlighting the need for targeted healthcare strategies focused on older adults. Advanced machine learning techniques, along with interpretative tools such as LIME (Local Interpretable Model-agnostic Explanations) and SHAP (Shapley Additive Explanations), were employed to analyze the influence of factors such as age, average glucose level, body mass index (BMI), hypertension, and heart disease on stroke predictions. Results showed that age emerged as a critical predictor, with individuals over 61 years exhibiting the highest risk, followed by elevated glucose levels and hypertension.Statistical analyses, including t-tests, demonstrated significant relationships between both BMI and glucose levels with stroke outcomes, supported by t-

statistics of 4.023 and p-values below 0.0001. Further correlation analyses revealed positive associations between age, BMI, and hypertension, reinforcing the multifaceted nature of stroke risk factors. These findings underscore the necessity of interpreting complex machine learning models to generate actionable insights for clinical practice, ultimately enhancing predictive accuracy and patient care strategies aimed at stroke prevention. Keywords Stroke, Hypertension, LIME (Local Interpretable Model-agnostic Explanations), SHAP (Shapley Additive Explanations), Body Mass Index (BMI), Glucose Levels

Keywords Stroke, Hypertension, LIME (Local Interpretable Model-agnostic Explanations), SHAP (Shapley Additive Explanations), Body Mass Index (BMI), Glucose Levels

DYNAMICAL SYSTEMS PERSPECTIVE OF PERCEPTUAL DECISION-MAKING AND METACOGNITION

Professor Dr. KongFatt Wong-Lin

Cognitive Neuroscience and Neurotechnology, Intelligent Systems Research Centre, School of Computing, Engineering, and Intelligent Systems, Ulster University, UK

Abstract/Synopsis: Perceptual decision-making relies on sensory information to guide the selection of a single choice. Extensive research in cognitive and systems neuroscience has elucidated its neural underpinnings by examining activity across individual neurons, neuronal populations, and large-scale brain networks. Moreover, each decision involves a certain degree of uncertainty—or confidence—that represents a metacognitive dimension of the process. In this talk, I will discuss the development of theoretical models that capture the dynamics of neural activity and decision-making behaviour, and show how dynamical systems theory can reveal the underlying mechanisms and computational principles. I will also highlight how these models may relate to classical probabilistic approaches and emerging quantum decision-theoretic framework.

COMPUTATIONAL APPROACHES IN PRODRUG DISCOVERY: A BIOINFORMATICS PERSPECTIVE

Dr. Dhanalakshmi G

Department of Biochemistry, Padmashree Institute of management and sciences, Bangalore, Karnataka, India

Abstract

In drug development, prodrug design is a crucial strategy aimed at improving the pharmacokinetic and pharmacodynamic properties of active drugs. Bioinformatics plays a vital role in modern prodrug discovery by providing computational tools to predict, optimize, and validate prodrug characteristics before experimental testing. The process begins with target identification, using databases like Gene Ontology and UniProt to identify relevant biological targets and enzymes involved in the activation of prodrugs. Once the target is identified, computational chemistry tools such as pharmacophore design, QSAR modeling, and molecular docking are used to design prodrug candidates that enhance properties like solubility, stability, and bioavailability.

A critical component of prodrug design is the prediction of metabolism. Tools such as MetaSite and StarDrop model metabolic pathways and predict interactions with key enzymes, such as CYP450 and esterases, to ensure the prodrug is effectively converted to its active form. ADME-Tox profiling is essential for evaluating the pharmacokinetic properties and toxicity of prodrugs, with tools like SwissADME and ADMETlab assessing aspects such as absorption, distribution, metabolism, excretion, and toxicity. Bioinformatics also helps assess potential drug-drug interactions using resources like DrugBank, ensuring that prodrugs are safe for use in combination with other medications. Moreover, bioinformatics facilitates the design of targeted prodrugs, which are activated selectively by enzymes overexpressed in specific tissues (e.g., cancer cells), reducing systemic toxicity. The iterative process of validation is guided by bioinformatics models to ensure that prodrug candidates meet the desired pharmacological profiles before moving to experimental testing. In summary, bioinformatics accelerates the discovery and optimization of prodrugs, enabling the development of safer, more effective drugs with enhanced bioavailability, targeted action, and reduced side effects.

PRESENTATION ABSTRACTS

FRUIT PEEL BIOFERTILIZER: HARNESSING SPOILAGE FOR SUSTAINABLE AGRICULTURE

Omkar Baliram Agale,

Institute of Biosciences and Technology, MGM University, Chh. Sambhajinagar

***Abstract*:** The study explores the potential of using fruit peel wastes, specifically from papaya and watermelon, as a biofertilizer through microbial fermentation. Fruit peels were subjected to spoilage over a period of two weeks to facilitate microbial growth. Microbial isolates were then obtained through a series of serial dilutions, followed by Gram staining and biochemical analysis, including the iodine test, to characterize and confirm the presence of amylase-producing bacteria. The significance of the iodine test lies in its ability to identify starch-degrading microorganisms, which are crucial for nutrient cycling in the soil. The biofertilizer derived from the spoiled fruit peels was then applied to fenugreek (*Trigonella foenum-graecum*) plants for a period of 20 days. A comparative analysis was conducted between plants treated with the biofertilizer and a control group with no fertilizer application. Plant growth rates were meticulously measured and compared to evaluate the efficacy of the biofertilizer.

The results indicated a significant enhancement in plant growth in the biofertilizer-treated group, demonstrating the potential of fruit peel waste as an effective and sustainable biofertilizer. This study highlights the dual benefits of waste valorisation and improved agricultural productivity, paving the way for eco-friendly farming practices.

COMPARATIVE ANTIBACTERIAL ANALYSIS OF NATURAL PRODUCTS AGAINST STAPHYLOCOCCUS AUREUS ISOLATED FROM FRESH WOUND SAMPLES.

Akash Ajay Choudhari

Institute of Biosciences and Technology, MGM University, Chh. Sambhajinagar

*Abstrac*t: This study investigates the antibacterial efficacy of natural products, namely honey, Aloe vera, Neem, and Turmeric, against Staphylococcus aureus strains isolated from fresh wound samples collected via swab tubes by medical professionals. The isolated bacteria were cultured on Mannitol Salt Agar and incubated at 36°C for 24-48 hours to facilitate growth and verify their identity through characteristic morphology and biochemical tests. Staphylococcus aureus growth was confirmed, and subsequent antibacterial testing was performed using the well diffusion method to assess the inhibitory effects of the natural products.

The well diffusion technique provided quantitative data on the zones of inhibition surrounding wells containing various concentrations of each natural product. Comparative analysis of these results highlighted variations in antibacterial efficacy, elucidating potential differences in therapeutic effectiveness among the tested agents. This research contributes valuable insights into the utilization of natural products as alternative or adjunctive treatments against Staphylococcus aureus infections, offering implications for future therapeutic strategies and the development of novel antimicrobial agents.

OPTIMIZING TESTOSTERONE THROUGH NUTRITION: A COMPREHENSIVE REVIEW.

Ronit Rohidas Savale

Institute of Biosciences and Technology, MGM University, Chh. Sambhajinagar

Abstract: Testosterone, a pivotal hormone in male physiology, plays a crucial role in regulating various physiological processes, including sex differentiation, spermatogenesis, and the development of secondary sexual characteristics. The modulation of testosterone levels in young adults is influenced by a myriad of factors, among which dietary habits and lifestyle choices stand prominent. This scientific review delves into the intricate relationship between nutrition, particularly macronutrient composition, and circulating testosterone levels, with a focus on its implications for athletic performance. We examine recent research findings, encompassing systematic reviews, meta-analyses, and intervention studies, to elucidate the effects of dietary factors on testosterone synthesis, metabolism, and bioavailability. Specifically, we explore the impact of dietary fat quality, protein intake, caffeine consumption, and carbonated beverage consumption on testosterone levels in men, drawing insights from both observational studies and randomized controlled trials. Additionally, we investigate the potential role of soy consumption and junk food intake in modulating testosterone levels and discuss the underlying mechanisms and clinical implications. Furthermore, we underscore the need for well-designed studies with larger sample sizes and rigorous methodologies to provide conclusive evidence regarding the effects of specific nutrients on testosterone levels. Ultimately, this review aims to enhance our understanding of the complex interplay between diet, hormones, and athletic performance, paving the way for the development of tailored nutritional strategies to optimize testosterone levels and promote overall health and athletic success in young adults.

EXPLORING THE INTERSECTION OF PSYCHOLOGY, PARAPSYCHOLOGY, AND PHILOSOPHY

K.S. Sri Rajeshwari Devi

Counselling psychologist, Author, Life skill Coach and Grapho-Analytical Therapist

srirajikaa@gmail.com

Abstract:

This paper examines the interconnected yet distinct fields of psychology, parapsychology, and philosophy, each offering unique perspectives on the complexities of human cognition, behavior, and consciousness. Psychology, rooted in empirical methods, investigates human behavior, mental processes, and social interactions through rigorous scientific inquiry. Parapsychology, controversial yet intriguing, explores phenomena beyond conventional understanding, such as ESP and psychokinesis, using scientific methods to probe claims of paranormal abilities. Philosophy, meanwhile, delves into fundamental questions about existence, knowledge, ethics, and the mind, employing rational inquiry to scrutinize the nature of reality and human experience. This abstract encapsulates how these disciplines intersect and diverge, contributing collectively to our understanding of the human condition and the mysteries of consciousness.

Key words: Consciousness, Telepathy, Intuition, paranormal phenomenon, and Extra Sensory Perception (ESP)

IDENTIFICATION AND CHARACTERIZATION OF THE TAU PROTEIN MUTATION AND THEIR ROLE IN PROTEIN STRUCTURE AND FUNCTION

Sreekanya Roy[1,2,#], Sima Biswas[1], Rakhi Dasgupta[1,*]and Angshuman Bagchi[1,2,*]

[1]Department of Biochemistry and Biophysics, University of Kalyani, Kalyani, Nadia, West Bengal, India

[2]BIF Centre, University of Kalyani, Kalyani, Nadia, West Bengal, India

#Presenting author e-mail ID: sreekanyabcbp23@klyuniv.ac.in

*Corresponding author e-mail ID: RDG: rdgadg@gmail.com

AB: angshu@klyuniv.ac.in

ABSTRACT

The increase in global life expectancy rate has made neurodegenerative diseases to be more alarming. Alzheimer's disease is the most prevalent among them. There are a number of theories that are being proposed to uncover this disease pathogenesis; Tau pathology is one of the more accepted ones along with Amyloid beta hypothesis. The key protein being involved in this hypothesis is the tau protein and hyperphosphorylation of this protein leads to formation of neurofibrillary tangles. These neurofibrillary tangles are indeed one of the critical clinical hallmarks of Alzheimer's disease. Several researchers identified clinical variations (SNPs) that are affecting the tau protein. Yet these studies fail to predict the effect of these SNPs on tau protein's structure and function as well as their damaging propensity for the disease. We first collected the mutations/SNPs which are located in the Tau protein from literature. In this study, we tried to predict the effect of these SNPs and categorised them based on their pathogenicity score. We used PredictSNP tool which is a combination tool of several in silico mutation prediction tools. Therefore, based on the prediction score we scored the 32 selected mutants according to their probability of being damaging for individuals. Moreover, we also categorized these pathogenic mutations on the basis of their damaging effect and identified the most damaging SNP present in the tau protein. Therefore, these insights will help us to get a clear understanding of the disease caused by these SNPs.

Keywords: Alzheimer's Disease; Tau pathology; tau protein; SNPs,

CIRCULAR RNA: AN IMPORTANT BIOMOLECULE TO CONTROL GASTRIC CANCER

Jit Mondal[1,#], Anirban Nandy[1], Sreekanya Roy[1,2], Sima Biswas[1], Dipanjan Guha[2] and Angshuman Bagchi[1,*]

[1]Department of Biochemistry and Biophysics, University of Kalyani, Kalyani, Nadia, West Bengal, India

[2]BIF Centre, University of Kalyani, Kalyani, Nadia, West Bengal, India

Presenting author E-mail ID: jitbcbp23@klyuniv.ac.in

* Corresponding author E-mail IDs: angshumanb@gmail.com, angshu@klyuniv.ac.in

Abstract

Recent studies with circRNAs indicate that, circRNAs can act as importantbiomarkersfor differentlethaldiseases including cancer.The involvement of severalRNAs, including circRNAswith cancer, is well established. Nevertheless, the underlying mechanism and regulation of many circRNAs are yet to be established. Our group has previously reportedthe unique link between circMAPK1 and circPGD, thecircRNAs,which were found to be associated with gastric cancer. CircPGDwas found to be the positive regulator and circMAPK1 is the negative regulatorof gastric cancer.Unlike circMAPK1 and circPGD, circDLG1 was also found to have progressive impacts on gastric cancer.Based on the evidence from literature search, we identified these three circRNAs that have their abilities to control gastric cancer and they exert their roles by sponging certain miRNAs. The preliminary information about those three circRNAs was obtained from circBase and circBank databases.The information about thecommon miRNAs associated with those three circRNAswas obtained from miRDB database. To corroborate the previous findings,information from circRNADisease database v2.0 was used.The objective of this study was to predictthe mechanism of action of these circular RNAs and also to report the associated pathways (if any), which are related to cancer, emphasizing on gastric cancer.The tool DAVID was used in this regard.miRNA (mir-141-3p) was found to be significant and have sponging site on circDLG1.Prediction on thefunctional genes associated with the miRNAs and these three circRNAs were also investigated, andit was

observed that they can influence numerous physiological processes and malignancies, including gastric cancer. In a nut-shell, the study reports aboutthe association of the three important circRNAs with various cancer pathways, including gastric cancer, which might be useful in future drug discovery process.

Keywords: CircRNA; miRNA; Online databases; Gastric cancer; Pathway prediction

SAFED MUSLI (*CHLOROPHYTUM BORIVILIANUM L.*) TUBER MEDIATED GREEN SYNTHESIS OF SILVER NANOPARTICLES AND THEIR CHARACTERIZATION

Rupali Taur[1], Shantanu Chavan[2], Avishkar Chonde[1], Kirti Sapkal[1] and Pallavi More[1]

[1]Institute of Biosciences and Technology, MGMU, Aurangabad, Maharashtra, India

[2]Sant Dnyaneshwar Mahavidyala Soyagano,Maharashtra, India.

Corresponding Author- Dr. Rupali Taur (Asst Prof.) Institute of Biosciences and Technology, MGMU, Aurangabad, Maharashtra, India

Synthesis of silver nanoparticles using plant extracts is one of the most rapid, efficient, economical, and ecofriendly strategy which diminish the use of toxic chemicals. However in current years, various environment friendly techniques for the efficient synthesis of silver nanoparticles have been demonstrated by utilizing aqueous extracts of plant parts, This investigation revealed the green synthesis of silver nanoparticles (AgNPs) using root extracts *Chlorophytum borivilianum*. AgNO3 solution used with the aqueous root extract revealed a change in color from yellow to brown showing to the bioreduction reaction. extracted AgNPs were characterized by using Scanning electron microscopy (SEM), X-ray Diffraction (XRD), and Fourier Transform Infrared Spectroscopy (FTIR). XRD pattern with typical peaks indicated the crystalline nature of silver. SEM analysis confirmed the presence of spherical-shaped AgNPs, having an average size of 48.0 nm. The FTIR analysis revealed the nanoparticles were stabilized by non-aromatic compounds present in aqueous extract and its significant role in the formation of bioreduction to form nanoparticles. Hence, these silver nanoparticles (AgNPs) possess significant power for utilization in various biomedical applications in coming era.

Keywords: *Chlorophytum borivilianum,* aquous extract, AgNPs, bioreduction.

DEVELOPING RESILIENT SPEECH EMOTION RECOGNITION SYSTEMS THROUGH DEEP LEARNING AND AUDIO AUGMENTATION FOR ENHANCED EMOTION DETECTION

Irfan Chaugule

MGM University, Dr. G.Y. Pathrikar College Of Computer Science And Information Technology, N-6, M G M, Chhatrapati Sambhajinagar (Aurangabad), Maharashtra 431003

Email ID: irfanchaugule@gmail.com

Abstract:

Speech Emotion Recognition (SER) has emerged as a critical area in human-computer interaction, aiming to enable systems to recognize and respond to human emotions expressed through speech. This research focuses on utilizing deep learning techniques to advance the performance of SER systems, particularly in noisy and variable conditions. We present a comprehensive approach, starting with the preparation of audio datasets, followed by the application of various augmentation techniques such as Gaussian noise, pitch shifting, time stretching, and time shifting, aimed at simulating real-world distortions. These augmentations, implemented using the *audiomentations* library, enhance the robustness of machine learning models by diversifying the training data.

We further explore the efficacy of deep learning models, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), in recognizing emotional states across different speech patterns. Initial results demonstrate significant improvements in model generalization, particularly in handling diverse audio conditions. This study contributes to the growing body of work on SER by improving model robustness through data augmentation, with promising results that lay the groundwork for more adaptive and emotion-aware systems.

MACHINE LEARNING DRIVEN SUBTYPE AND MUTATION CLASSIFICATION IN GLIOMAS USING TRANSCRIPTOMIC DATA

Shreeya Pahune[1] and Bhaswar Ghosh [2,*]

[1,2] Center for Computational Natural Sciences and Bioinformatics, International Institute of Information Technology, Hyderabad 500032, Telangana, India

[1] Undergraduate Student. Email: shreeya.pahune@research.iiit.ac.in (+919381859976)

[2] Assistant Professor. Email: bhaswar.ghosh@iiit.ac.in (+918335037321)

Abstract

Gliomas, accounting for nearly one-third of all brain tumours, present themselves as significant medical challenges owing to their heterogeneous and aggressive nature. Given that traditional diagnostic techniques often fall short of capturing the entire genetic landscape, there are minimal options currently present in personalised medicine, thus adding to existing challenges. In order to aid precision oncology, we present a machine learning (ML) framework for glioma subtype classification and mutation prediction using RNA sequencing data. We used a two-step feature selection procedure that included XGBoost and LASSO, leveraging RNA sequencing and somatic mutation data of 667 glioma samples from the Cancer Genome Atlas (TCGA) Pan-Cancer project to identify key genes that are associated with glioma subtypes and important mutations. Metrics such as AUROC, F1-Score, and Matthews Correlation Coefficient (MCC) were used to train and evaluate the supervised machine learning models, including Random Forest and Stochastic Gradient Descent. Our framework performed robustly despite the class imbalance present in the dataset. We achieved a high classification accuracy for Diffuse Glioma (DG) subtypes (AUROC 0.90-1.00), distinguishing between low-grade gliomas (LGGs) and glioblastoma multiforme (GBM). Mutation status predictions for the key prognostic genes, including IDH, TP53, and ATRX, achieved high AUROC scores of 0.89-0.98 and MCC values of 0.72-0.97. Further, we used feature sets from our top-performing models to perform gene set enrichment as part of our post-analysis. This confirmed the biological significance of identified genes related to established carcinogenic pathways and further validated the clinical applicability of our technique. Our ML-driven framework presented offers a scalable, data-driven solution in computational oncology with potential applications across cancer types.

UNRAVELLING PLASMODIUM FALCIPARUM'S LIFE CYCLE: INSIGHTS FROM INTEGRATED TRANSCRIPTOMIC AND MACHINE LEARNING APPROACHES

Shreeya Pahune[1] and Bhaswar Ghosh [2,*]

[1,2] Center for Computational Natural Sciences and Bioinformatics, International Institute of Information Technology, Hyderabad 500032, Telangana, India

[1] Graduate Student. Email: shreeya.pahune@research.iiit.ac.in (+919381859976)

[2] Assistant Professor. Email: bhaswar.ghosh@iiit.ac.in (+918335037321)

Abstract

This study uses a transcriptomic and machine learning-based approach to gain a comprehensive understanding of the genomic landscape of Plasmodium falciparum, a malaria-causing parasite. This enables us to understand this life-threatening disease better and thus identify new treatment targets and intervention methods. We used single-cell RNA sequencing data from the Malarial Cell Atlas with 5,177 genes across 37,624 cells. Our approach combined static analyses, including highly variable gene identification, differential expression analysis, and M3Drop (a feature selection method based on noise filtration), to obtain stage-specific markers along with trajectory analysis (Monocle3) to identify top genes active during stage progression. The feature selection analyses were done for each stage and lifecycle gene, which are genes that remain active throughout the lifecycle. These complementary approaches gave us distinct gene sets: static analysis revealed genes necessary for stage-specific functions, while dynamic trajectory analysis highlighted genes that play an important role in cellular development. A neural network classifier trained on these gene sets achieved high accuracy (Accuracy: 96%; F1-score: 0.96) using trajectory-derived genes, performing better than the model trained on the entire gene set, indicating that the gene set has captured stage-specific signatures. Further, to validate the biological significance of the shortlisted genes, we performed a Gene Ontology analysis. We found that our analysis revealed results that are in line with existing literature. The ring stage genes were linked to immune evasion, trophozoite genes to haemoglobin digestion, schizont genes to merozoite invasion, and gametocyte genes to sexual differentiation. We plan to build on this work by studying the behaviour of important genes using a Poisson-Beta model to infer the kinetics of key genes, which will help us better understand how they change throughout the parasite's life cycle.

PREDICTION OF PARKINSON'S DISEASE USING MULTIMODAL MACHINE LEARNING AND INTEGRATION TECHNIQUES

Shreeya Pahune[1], Kushagra Agarwal[2], Bhaswar Ghosh[1], Nita Parekh[1,*]

[1] Center for Computational Natural Sciences and Bioinformatics, International Institute of Information Technology, Hyderabad 500032, Telangana, India

[2] School of Computer Science, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA

Shreeya Pahune: Graduate Student, IIIT-H, (shreeya.pahune@research.iiit.ac.in) (+91 9381859976)

Kushagra Agarwal: Graduate Student, Carnegie Mellon University, (kagarwa2@andrew.cmu.edu) (+1 412-589-2871)

Bhaswar Ghosh: Assistant Professor, IIIT-H (bhaswar.ghosh@iiit.ac.in) (+91 8335037321)

Nita Parekh: Associate Professor, IIIT-H (nita@iiit.ac.in) (+91 4066531000)

Abstract

Parkinson's Disease (PD) is a serious neurodegenerative disorder for which reliable early diagnostic tests are lacking. This study presents a method to predict PD using multimodal data from the Parkinson Progression Marker Initiative (PPMI), integrating clinico-demographic, biospecimen, and genetic data to enhance predictive accuracy and support timely interventions. We analyzed data from 598 participants (171 healthy controls and 427 PD patients) across three data types: 29 clinical features, five cerebrospinal fluid biomarkers, and 154 SNPs (single nucleotide polymorphisms), selected through a biology-driven feature selection method. Three multimodal integration strategies—early, intermediate, and late—were employed to train machine learning models, including LightGBM and Multilayer Perceptron (MLP), each optimized with hyperparameter tuning and cross-validation. In the early integration strategy, we combined features from all data types into a single set, leveraging complementary information to boost predictive accuracy. The intermediate integration used autoencoders to reduce features to a 12-dimensional vector, which was then input into a Neural Network classifier. Late integration combined

outputs from top models for each data type using ensemble methods like Voting Classifier and Stacking. Among these, early integration yielded the highest performance, with Support Vector Machines achieving 90% accuracy, 0.98 AUC-ROC, 0.99 precision, and 0.93 F1 score. The intermediate integration with the Neural Network classifier followed, showing an AUC-ROC of 0.84 and an F1 score of 0.85. Feature importance analysis highlighted two clinical scores—UPSIT (related to sensory decline) and SCOPA (measuring autonomic dysfunction)—and two SNPs (chr4_90755939_A_G and chr4_90646886_G_A, both linked to PD susceptibility) as key predictors, underscoring their relevance in PD diagnosis. Integrating multiple data types facilitates early diagnosis and treatment of PD and significantly improves disease prediction accuracy, providing a comprehensive understanding of the disease's complexity.

NETWORK-DRIVEN DRUG REPURPOSING: INSIGHTS FROM GRAPH NEURAL NETWORKS AND RECOMMENDATION MODELS

Kushagra Agarwal [1,*], Shreeya Pahune [2], Hemant Chandak [3], Sumeet Agrawal [4]

[1] Graduate Student, Carnegie Mellon University (kagarwa2@andrew.cmu.edu) (+1 412-589-2871)

[2] Graduate Student, IIIT-Hyderabad (shreeya.pahune@research.iiit.ac.in) (+91 9381859976)

[3] Student, Manipal University Jaipur (hemant.209303369@muj.manipal.edu) (+91 7890978539)

[4] Technical Lead, Avalara (sumeetagrawal.srm@gmail.com) (+91 9891487397)

Correspondence: Kushagra Agarwal, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213, USA

Abstract

The rising cost of drug discovery, coupled with a stagnation in the approval of novel treatments, highlights the urgent need for innovative strategies such as drug repurposing. Pharmaceutical companies invest roughly 10-15 years and $2.6 billion to get a single FDA-approved drug to market. The COVID-19 pandemic further underscored the necessity of quickly identifying existing drugs with potential efficacy against a fast-spreading virus to curtail the pandemic. In this study, we perform a comparative analysis of several Graph Neural Networks (GNNs) and recommendation system models to address drug repurposing. We construct an integrated graph that combines Protein-Protein Interaction networks, Drug-Target Protein graphs, Disease-Protein associations, and Drug-Disease links. We leverage a network learning paradigm implemented over this complex graph via both node-agnostic and heterogeneous graph techniques for link prediction in drug-disease pairs. We implement a Heterogeneous Graph Transformer (HGT) model that processes three node types (drugs, diseases, proteins) and four edge types. The HGT achieved an AUC-ROC of 0.985 and an F1-score of 0.90, demonstrating its efficacy in predicting drug repurposing candidates. Additionally, we compared several node-agnostic GNN architectures, including Graph Convolutional Networks, Graph Attention Networks, GraphSAGE, and Graph Isomorphism Networks. All architectures performed comparably, with an AUC-ROC of

around 0.98. However, when framing the drug repurposing task as a recommendation problem using Matrix Factorization with side information, we observed a significant drop in performance, with the AUC-ROC falling to 0.82. This performance degradation highlights the importance of incorporating Protein-Protein Interaction networks in the modeling process, as matrix factorization fails to capture these complex network effects critical for drug repurposing. Our models ranked 6,158 drugs based on their predicted efficacy in treating COVID-19, providing a valuable tool for prioritizing clinical trials and further research. Beyond COVID-19, such an integrated framework can allow us to uncover drug-repurposing prospects for any other novel diseases in a significantly more efficient and cost-effective way.

Code Availability: https://github.com/kushagragarwal2443/Drug_Repurposing_GNN

TOWARDS EFFECTIVE MCQ GENERATION: SEMANTIC AND MACHINE LEARNING APPROACHES.

Jeevan Tonde[1] and Satish Sankaye[2]

[1]Ph.D. Scholar, Dr. G.Y. Pathrikar College of Computer Science and Information Technology, MGM University, Chhatrapati Sambhajinagar

Assistant Professor, MIT Arts, Commerce & Science College, Alandi(D)

[2]Associate Professor, Dr. G.Y. Pathrikar College of Computer Science and Information Technology, MGM University, Chhatrapati Sambhajinagar

E-mail: [1]jptonde@gmail.com, [2]sankayesr@gmail.com.

Abstract

Multiple Choice Questions (MCQs) are widely used as an assessment tool due to their scalability, automated evaluation, and flexibility across educational and industrial domains. However, manually creating MCQs is labor-intensive, requiring expertise to generate the question stem, correct answer ("key"), and plausible distractors. Recent advancements in Automatic Question Generation (AQG) and Natural Language Processing (NLP) have facilitated the automation of this process. This paper presents a comparative study of two MCQ generation techniques: a hybrid semantic-rule-based approach, and a machine-learning-based approach. The semantic-rule-based method combines linguistic and conceptual understanding to ensure that the generated questions are both meaningful and contextually appropriate, while rule-based algorithms ensure the questions conform to grammatical and structural standards. In contrast, the machine-learning-based approach uses advanced models like neural networks and transformers to generate both questions and distractors by identifying patterns from large datasets To support this research, we prepared a comprehensive dataset consists of well-structured Multiple Choice Questions (MCQs), each containing a question stem, a correct answer, and several plausible distractors. The questions were carefully curated from reliable educational resources and peer-reviewed materials to ensure accuracy and relevance. The generated MCQs are then evaluated using various metrics, including BLEU, ROUGE and human assessment. This research offers valuable insights for improving automated MCQ generation and its application in intelligent tutoring systems and e-learning platforms.

ENHANCING DATA SECURITY WITH DNA METHYLATION-BASED ENCRYPTION TECHNIQUES

Rajat Pandey, Tejal Dube, Manojkumar Charandas Ramteke

Indian Institute of Technology Delhi, IIT Delhi Main Rd, IIT Campus, Hauz Khas, New Delhi, Delhi 110016

Contact info: Rajat Pandey (+91 9674893920); **Mail Id:** chz228172@iitd.ac.in

Abstract:

In an era of exponential data growth and increasing cyber threats, ensuring robust data security has become imperative. This paper presents a novel DNA methylation + RSA-based hybrid algorithm for text encryption, leveraging the complexity of biological processes and cryptographic techniques to enhance data security. By simulating DNA methylation—a process where methyl groups are added to cytosine bases in DNA sequences—an additional layer of encryption is introduced, significantly increasing resistance to brute-force, statistical, and differential attacks. The algorithm encodes text data into DNA sequences, applies methylation to key nucleotides, and combines it with RSA encryption for enhanced cryptographic strength.

Extensive evaluations demonstrate the effectiveness of the technique. High randomness is achieved in ciphertext, validated by statistical tests, reflecting the unpredictability of the encrypted data. Encryption and decryption processes maintain high accuracy, with minimal computational overhead and high scalability across varying data sizes and key lengths. Decryption is achieved successfully, confirming data integrity and fidelity in reconstruction.

The findings suggest that the DNA methylation + RSA hybrid encryption algorithm offers a secure, bio-inspired cryptographic solution for sensitive data transmission, with potential applications in domains requiring high confidentiality and robustness, such as healthcare, defence, and secure communications.

Keywords: DNA methylation, RSA, text encryption, bio-inspired cryptography, data security, hybrid encryption algorithms.

BIG DATA IN BIOINFORMATICS

Sourabh Maid, Rudra Verma, Nachiket Dhamane

BSc Bioinformatics First Year,

Deogiri College, Chhatrapati Sambhajinagar

Email: nachiketdhamane687@gmail

Phone no: 8849770117/9529510413

Abstract: Bioinformatics integrates various fields to manage and analyse large biological data. With the surge in data from genome-wide experiments and electronic records, advanced computing and updated educational methods are crucial. The Human Genome Project has improved data collection and reduced sequencing costs. Tools like BLAST and FASTA are key for sequence comparison, while bioinformatics also studies gene effects on traits via protein and metabolite analysis. The shift from CPUs to GPUs has accelerated research, and methods like SOM, ANN, SVM, fuzzy logic, and neuro-fuzzy networks enhance data management and support cancer screening and prognosis. Key resources include NCBI (with GenBank and PubMed), DDBJ (for nucleotide sequences), and EMBL-EBI (offering Ensembl and UniProt).

Keyword

Biological Big Data, Bioinformatics, Analysis Tools.

INTERACTIONS BETWEEN LONG NON-CODING RNAS AND PROTEINS TO ELUCIDATE THE MECHANISM OF NEURODEGENERATIVE DISEASE ONSET

Aniruddha Biswas[1], Sima Biswas[2], Angshuman Bagchi[3]*

[1]JIS College of Engineering, Block-A, Phase-III, Kalyani, Nadia-741235, West Bengal, India, +91 8697357415, aniruddha.biswas@jiscollege.ac.in

[2,3]Department of Biochemistry and Biophysics, University of Kalyani, Kalyani, Nadia-741235, West Bengal, India, angshu@klyuniv.ac.in

*Corresponding author

Abstract

Neurodegenerative diseases are characterized by the progressive loss of nerve cells or myelin in different parts of the brain. Long non-coding RNAs (lncRNAs) are important players in neurodegeneration. LncRNAs, consisting of 200 nucleotides or more, perform a varied range of functions. They are involved in disease pathogenesis via mechanisms like acting as decoys, scaffolds, and miRNA sequestrators. These mechanisms include mitochondrial dysfunction, oxidative stress, and inflammation, which promote neurodegenerative disease progression. lncRNAs, through their protein-binding capabilities, influence gene expression, protein homeostasis, and cellular stress responses, which are crucial in neurodegenerative diseases. Less than 2% of the human genome's transcripts encode proteins, with the rest being ncRNAs, including lncRNAs. Studies show that there are over 19,000 lncRNAs in humans, playing essential roles in neurogenesis and CNS processes like metabolism and apoptosis. Key characteristics of lncRNAs include poor sequence conservation and tissue-specific expression patterns. LncRNAs have regulatory roles in gene expression by interacting with chromatin and proteins, forming complexes that control downstream genes. In neurodegenerative diseases, they interact with various proteins, influencing processes such as Aβ deposition in Alzheimer's disease and UPS dysfunction in Parkinson's disease. They can be categorized into five types based on genome location, including long intergenic and antisense lncRNAs. Examples include mutant huntingtin aggregates in Huntington's disease and TDP43 proteinopathies in ALS. Understanding lncRNA mechanisms will be useful to develop novel therapeutic strategies.

Keywords: *Neurodegenerative diseases (NDDs), long non- coding RNAs (lncRNAs), Alzheimer's disease (AD), Parkinson's disease (PD), Huntington's disease (HD), and amyotrophic lateral sclerosis (ALS), protein binding interaction.*

REVISITING SUSHRUTA SAMHITA AND INVESTIGATING THE POTENTIAL LINK BETWEEN ITS TREATMENT MODALITIES BY PREDICTING STUDY USING BIOINFORMATICS

Vaibhavi Jain[1*] and Arpita Kharat[1]

Department of Bioinformatics, MGM Institute of Biosciences & Technology, Chhatrapati Sambhaji-nagar- 431003 Maharashtra, India

Contact no: +91-8626035070, **Corresponding Author:** E-mail: jvaibhavi722@gmail.com

Abstract

This study delves into the therapeutic potential of compounds derived from the ancient Indian medical text, Sushruta Samhita, for treating diseases like tuberculosis, eye disorders, colic, leprosy, and chronic fever. By employing molecular docking simulations, we explored the interactions between bioactive compounds from these formulations and their corresponding disease-related gene receptors. Promising results were obtained for compounds like beta-carotene, tocopherols, vitamin A, phytosterols, and vitamin K in targeting specific disease pathways. This research bridges the gap between traditional Ayurvedic knowledge and modern drug discovery, offering a novel approach to developing effective and culturally relevant treatments.

Keywords: Ayurveda, Sushruta Samhita, molecular docking, drug discovery, tuberculosis, eye diseases, colic, leprosy, chronic fever, bioactive compounds, gene receptors, etc.

USAGE OF NATURAL LANGUAGE PROCESSING IN EXTRACTING THE INVALUABLE INSIGHTS FROM SUSHRUTA SAMHITA

Vaibhavi Jain[1]*, Arpita Kharat[1], Preenon Bagchi[2], and Shivani V[2]

[1]Department of Bioinformatics, MGM Institute of Biosciences & Technology; Chhatrapati Sambhajinagar - 431003 Maharashtra, India

[2] Karnataka Sanskrit University, Bengaluru, India

*Corresponding author, E-mail: jvaibhavi722@gmail.com

Abstract

Sushruta Samhita, an ancient Sanskrit text dating back to around 600 BCE, stands as a cornerstone in the history of medicine, particularly in the domain of surgery. Despite its age, the text contains a wealth of knowledge and insights that are still relevant in contemporary medical practice. In this study, we employ Natural Language Processing (NLP) techniques to extract invaluable insights from the Sushruta Samhita. By leveraging NLP methodologies such as text parsing, entity recognition, and semantic analysis, we delve into the intricate details of surgical techniques, anatomical descriptions, disease classifications, and therapeutic approaches outlined in the text. Through computational analysis, we aim to unearth hidden patterns, correlations, and practical wisdom embedded within the text, shedding light on the timeless wisdom of ancient Indian medicine. Our findings not only contribute to a deeper understanding of Sushruta's contributions to the field of medicine but also demonstrate the potential of NLP in unlocking the knowledge contained within ancient texts for contemporary scientific inquiry and medical practice.

Keywords: *Ayurveda, Natural Language Processing (NLP), Samhita Samhita, Knowledge Extraction, Traditional Medicine, Healthcare, Sanskrit, Ancient Wisdom, Google colab, Python*

EXPLORING SPIRITUAL WISDOM: NLP-BASED KNOWLEDGE EXTRACTION FROM THE WORKS OF SRI SRI THAKUR ANUKULCHANDRA

Arpita Kharat*[1], Vaibhavi Jain[1], Preenon Bagchi[2] and Shivani V[2]

[1]Department of Bioinformatics, MGM Institute of Biosciences & Technology; Chhatrapati Sambhajinagar - 431003 Maharashtra, India

[2]Dept. of Sastra, Karnataka Samskrit University, Bengaluru

*Corresponding author, E-mail: arpitakharat5188@gmail.com

ABSTRACT

Using Natural Language Processing (NLP) tools, an overview of Sri Sri Thakur Anukulchandra's dharma is presented in this abstract. The well-known spiritual teacher and philosopher Sri Sri Thakur Anukulchandra promoted holistic living by adhering to the values of love, service, and moral behaviour. In addition to promoting devotion, meditation, and the value of a guru-disciple relationship, his writings offer guidance to readers on spiritual development. The development of compassion and love for all people, with a particular emphasis on using service to others as a means of worship, are important themes. In addition, Thakur Anukulchandra explores family and social harmony, providing helpful guidance on forging enduring bonds and carrying out social obligations. His universalistic viewpoint encourages an inclusive spiritual approach by fostering respect for all religions and beliefs. He also stresses the value of knowledge, bettering oneself, and preserving one's physical and mental health. His teachings emphasise the value of moral behaviour and self-control while promoting a healthy lifestyle that integrates material and spiritual goals. The application of NLP techniques in this extraction reveals recurrent themes and key takeaways from the writings of Thakur Anukulchandra, offering a succinct summary of his extensive advice on leading a happy and satisfying life.

Keywords :- Knowledge Extraction, Natural Language Processing (NLP), Textual Analysis

IN SILICO STUDY OF *SYZYGIUM CUMINI* COMPONENTS WITH HMG-COA REDUCTASE FOR POTENTIAL ANTI-OBESITY EFFECTS

Navya Raj A R, Pruthivi A, Chandana S J and Dhanalakshmi G*

Department of Biochemistry, Padmashree Institute of Management and Sciences, Bangalore, Karnataka, India

*dhanu.bio@gmail.com

Syzygium cumini (Jamun), a tropical fruit renowned for its rich nutritional and medicinal properties, has been widely studied for its potential therapeutic applications. This in silico study investigates the bioactive phytochemicals of *Syzygium cumini* for their anticholesterol activity by analyzing their interactions with the target protein HMG-CoA reductase. Structures of selected phytochemicals, including flavonoids, anthocyanins, tannins, and essential oils, were retrieved from the PubChem database, prepared as ligands, and evaluated for their pharmacokinetic properties and Lipinski's Rule of Five.

Nine compounds, including Simvastatin (positive control), Dihydroquercetin, Gallic Acid, Ellagic Acid, Kaempferol, Quercetin, Mycaminose, and Tetradecanoic Acid, were identified as ligand candidates. Docking analysis revealed that Dihydroquercetin (-8.7), Kaempferol (-7.9), Ellagic Acid (-8.6), and Quercetin (-8.4) exhibited stronger binding affinities with HMG-CoA reductase than Simvastatin (-7.7), indicating their higher potency as anticholesterol agents.

These findings emphasize the therapeutic potential of Syzygium cumini phytochemicals, particularly for cholesterol management. Future in vitro and clinical studies are necessary to further validate these effects. It holds promise as an alternative plant-based medicine for obesity and related diseases.

Key words: Molecular Docking, Drug discovery, Syzygium cumini, HMG-CoA and Anti-Cholesterol.