

MINI PROJECT
AT
SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY

Submitted in partial fulfillment of the requirements for the award of a Master of Science
degree in Bioinformatics and Data Science

by

PRAVEEN RAJA K
REGISTER NO 43739040



DEPARTMENT OF BIOINFOMATICS AND
DATA SCIENCE

SCHOOL OF BIO CHEMICAL ENGINEERING

SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY (DEEMED TO
BE UNIVERSITY)

Accredited with Grade "A" by NAAC | 12B Status by UGC | Approved by AICTE
JEPPIAAR NAGAR, RAJIV GANDHI SALAI, CHENNAI - 600 119

April - 2024



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)

Accredited with "A" grade by NAAC | 12B Status by UGC | Approved by AICTE
Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai – 600 119

www.sathyabama.ac.in

DEPARTMENT OF BIOINFORMATICS AND DATASCIENCE

BONAFIDE CERTIFICATE

This is to certify that this Intern Report is the bonafide work of Praveen Raja K (43739040) who carried out an internship on "Mini Project" from September 2023 to November 2023.

Internal Guide

External Guide (if Applicable)

Head of the Department

Submitted for Viva-voce Examination held on: _____

Internal Examiner

External Examiner

DECLARATION

I, **PRAVEEN RAJA K** hereby declare that the Internship report entitled Mini Project done by me under the guidance of **Mrs. SUDHA** at Sathyabama Institute of Science and Technology is submitted in partial fulfillment of the requirements for the award of a **Master of Science in Bioinformatics and Data and Science**.

DATE:29/04/2024

PLACE: Chennai

SIGNATURE OF THE CANDIDATE

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to the Board of Management of SATHYABAMA for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

I convey my thanks to **Dr.REKHA CHAKRAVATHI M.E., PH.D. DEAN OF ARTS AND SCIENCE** and **Dr. Jemmy Christy Head of the department** for providing me necessary support and details at the right time during the progressive review.

I wish to express my thanks to all Teaching and Non-teaching staff members of the Department of Bioinformatics and Data Science who were helpful in many ways many Ways for the completion of Internship.

TABLE OF CONTENTS

S.NO	TITLE	PAGE
1	CYTOSCAPE	5
2	KM PLOTTER	11
3	BIOMARKER	12
4	DRUG TARGET	13
5	ABSTRACT	15
6	INTRODUCTION	16
7	METHODOLOGY	17
8	RESULT	18
9	DISCUSSION	19
10	DRUG DEVELOPMENT	21
11	MATLAB COURSE	22
12	REVIEW ARTICLE	33
13	PLAGARISM REPORT	35

CYTOSCAPE

AIM

To analyze gene expression data retrieved from the Gene Expression Omnibus (GEO) database and visualize the results in Cytoscape to elucidate the molecular mechanisms underlying biological processes or disease conditions.

MATERIALS

GEO2R
CYTOSCAPE
CYTOHUBBA
CLUEGO

GEO2R

GEO2R is a web-based tool provided by the Gene Expression Omnibus (GEO) that allows users to analyze gene expression data deposited in GEO datasets. It is designed to perform differential gene expression analysis, which helps identify genes that are differentially expressed between different experimental conditions or biological states.

CYTOSCAPE

Cytoscape is a powerful open-source software platform for visualizing, analyzing, and interpreting complex networks in biological systems. It is widely used by researchers in bioinformatics, systems biology, and computational biology to explore molecular interactions, pathways, and regulatory networks derived from various types of high-throughput data, such as gene expression, protein-protein interactions, and signaling pathways.

CYTOHUBBA

Cytoscape's CytoHubba is a plugin that offers various network analysis algorithms for identifying important nodes (e.g., proteins, genes) within biological networks. It focuses on finding nodes that are highly connected or have specific topological features, which can indicate their significance in the network's structure or function. Here is an overview of CytoHubba and its features.

CLUEGO

ClueGO is a Cytoscape plugin designed to analyze and visualize functional enrichment of gene ontology (GO) terms and biological pathways within biological networks. It is commonly used by researchers in bioinformatics and systems biology to gain insights into the biological significance of gene sets identified through high-throughput experiments, such as gene.

PROCEDURE

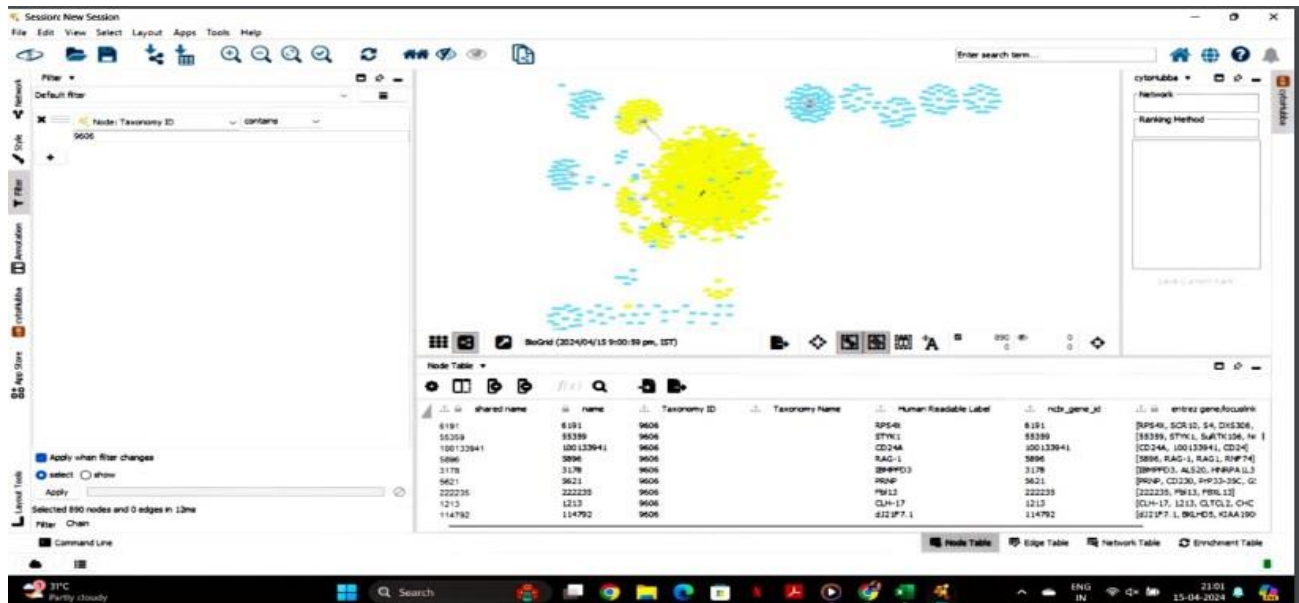
- For the database and disease pulmonary fibrosis Geo2R is the source of the geo dataset (GSE36701)
- After choosing the suitable parameters, the Geo dataset is retrieved.
- Genes that were upregulated genes were copied and paste into the cytoscapae network.
- Built a biological network using a BioGrid in cytohubba to analyses the network and identify importance node based various topologies algorithm to determine the the top 10 Genes ACAT-1, AAT3, ALS2CR7, AAC2, ALS10, ALK4, AD-006, ALK and ALS6
- Genes must be copied in cluego and continues cluego parameter such as species ontology and statistical settings.
- Run cluego analysis to identify functional categories or pathways associated with the selected nodes.

RESULT:

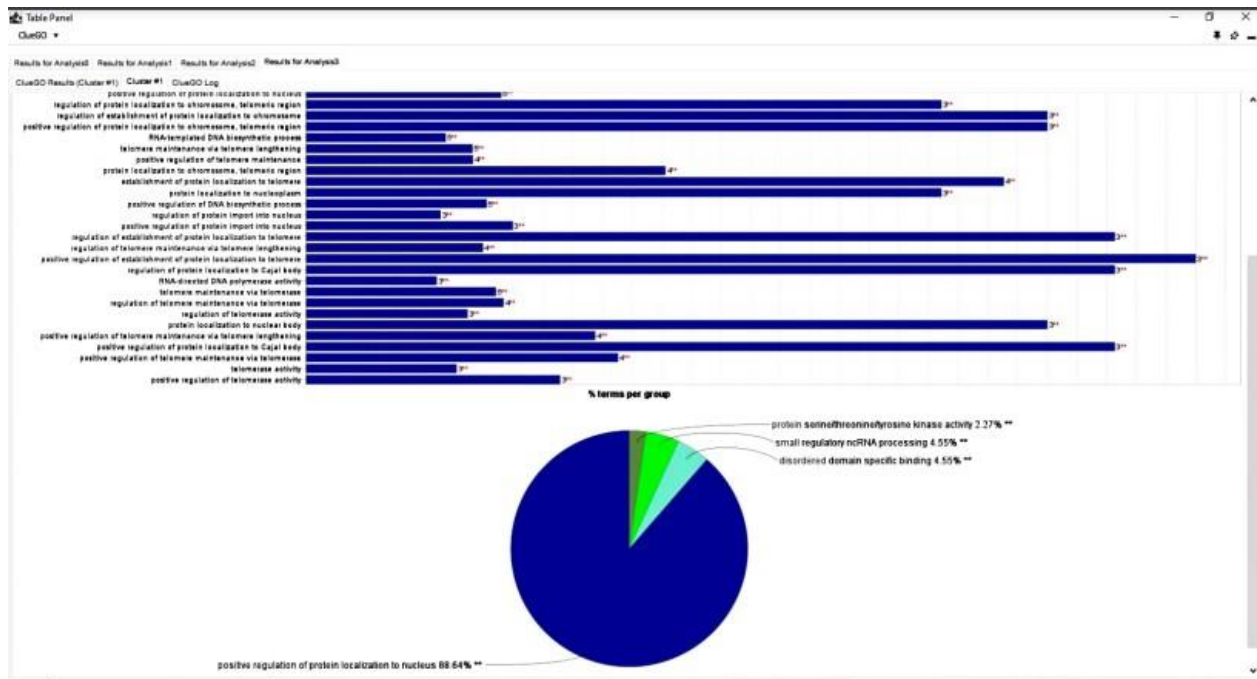
CYTOSCAPE



CYTOHUBBA



CLUEGO



INTERPRETATION

Interpreting the process of analyzing gene expression data from GEO and visualizing it in Cytoscape involves understanding its significance in molecular biology research. The integration of gene expression data analysis from GEO with visualization in Cytoscape offers a powerful.

CUMULATIVE SURVIVAL PROBABILITY CALCULATION USING KM PLOTTER

INTRODUCTION:

KM-plotter is the most sophisticated online survival analysis tool, performing all calculations in real time (not loading pre-calculated images). The background database is manually curated. Gene expression data and relapse free and overall survival information are downloaded from GEO, EGA and TCGA. The database is handled by a PostgreSQL server, which integrates gene expression and clinical data simultaneously. To analyze the prognostic value of a particular gene, the patient samples are split into two groups according to various quantile expressions of the proposed biomarker. The two patient cohorts are compared by a Kaplan-Meier survival plot, and the hazard ratio with 95% confidence intervals and log rank P value are calculated. Databases and clinical data are supervised and extended regularly.

AIM

Calculate Proposed gene ALK cumulative survival probabilities.

MATERIALS AND METHODS:

KM-PLOTTER: KM-plotter is the most sophisticated online survival analysis ,performing all calculations in real time (not loading pre-calculated images). The background database is manually curated. Gene expression data and relapse-free and overall survival information are downloaded from GEO, EGA and TCGA.

Procedure:

Kaplan Meier plotter home page was accessed using the following hyperlink

<https://kmplot.com/analysis/index.php?p=background>

Proposed gene/protein MX1 was submitted in the pan cancer portal

Below listed parameters were opted for analysis

Restriction subtype –All

Split patients by-Auto select best cut off

Survival: Overall survival

Survival: All

Restrict analysis based on cellular content: All

Sources for the database include GEO, EGA, TCGA, Metabric, Impact, and PubMed repositories.

INTERPRETATION

Bladder Carcinoma

The HR of PRB1 in bladder carcinoma is 1.42(1.05-1.93) 50% median survival between groups .
it represents the ratio of probability for low expression cohort (42.33) and high expression cohort (24.47)

BREAST CANCER

The HR of MX1 in breast cancer is 1.36(0.96 -1.93)50% median survival between groups. it represents the ratio of probability for low expression cohort (131.5) and high expression cohort (115.73)

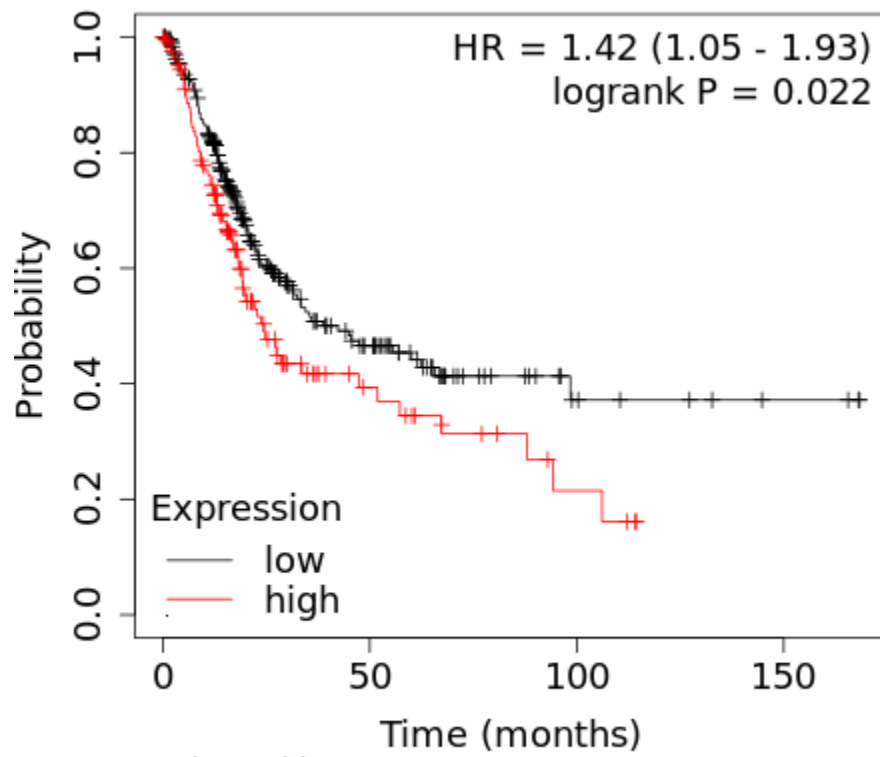
CERVICAL SQUAMOUS CELL CARCINOMA

The HR of MXI in cervical squamous cell carcinoma is 0.56(0.32-0.97)50% median survival between groups. it represents the ratio of probability for low expression cohort (NA)and high expression cohort (NA)

ESOPHAGAL SQUAMOUS CELL

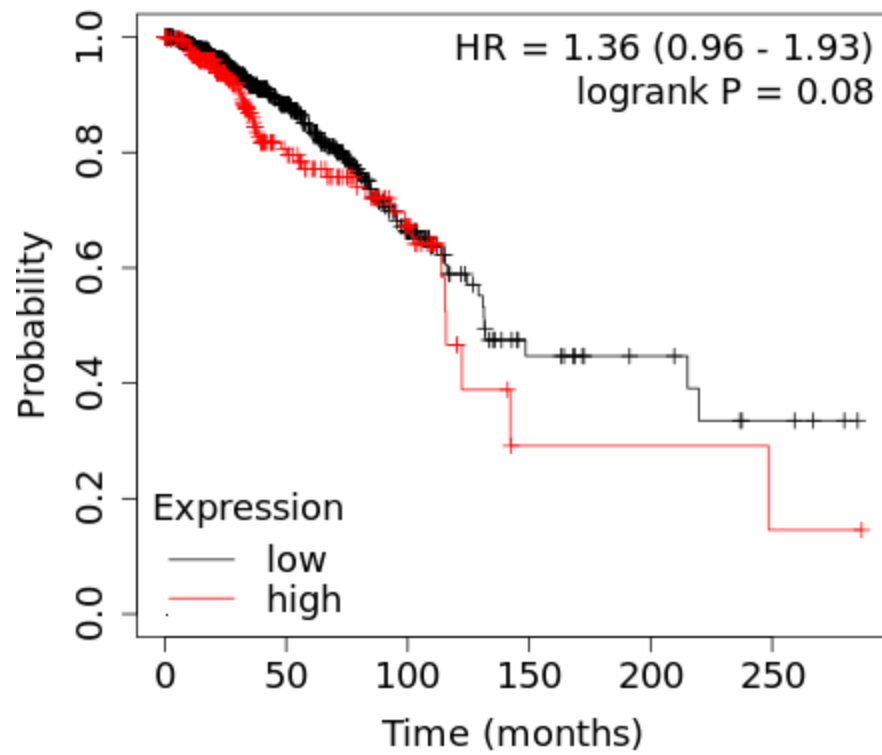
The HR of MXI in esophageal squamous cell is 0.52(0.28-0.99)50% median survival between groups. it represents the ratio of probability for low expression cohort (16) and high expression cohort (32.9)

ALK



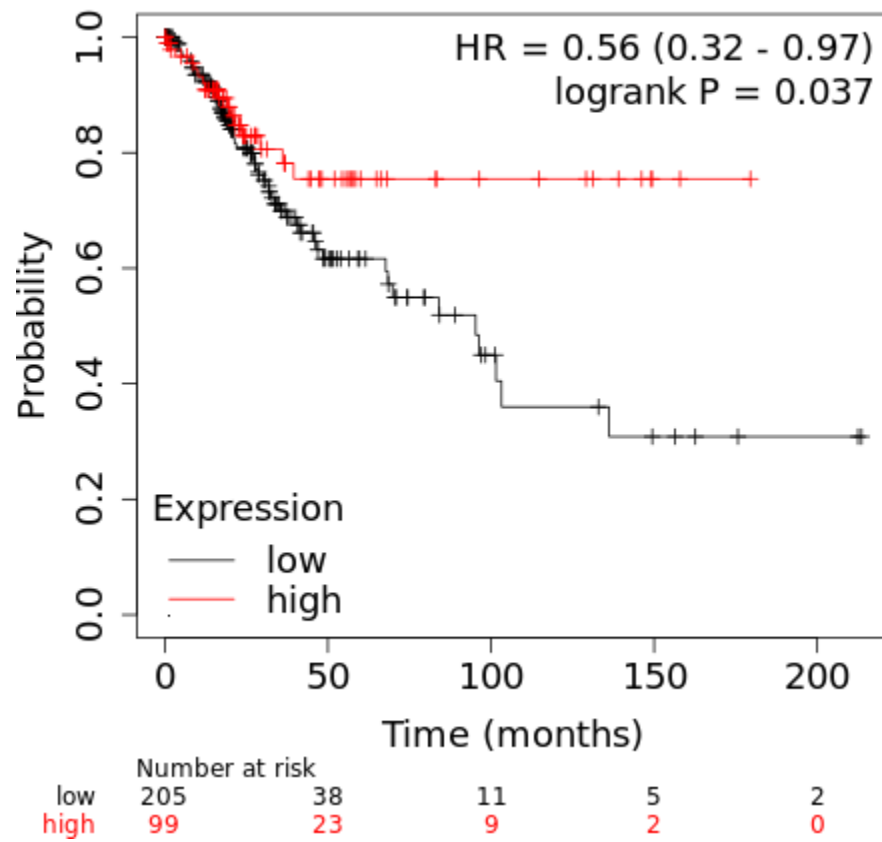
	Number at risk			
low	276	50	8	3
high	128	16	4	0

ALK

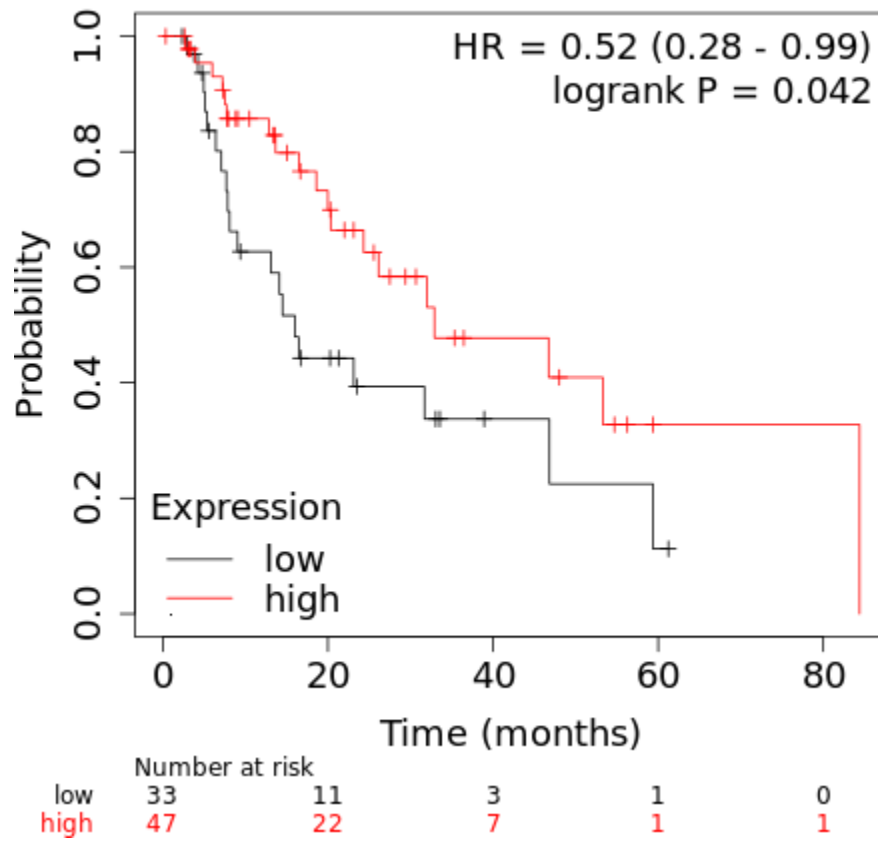


	Number at risk					
low	816	247	70	16	9	4
high	273	73	25	2	2	1

ALK



ALK



Biomarkers: identifying biological markers for disease Diagnosis and Therapeutic monitoring.

INTRODUCTION

The Marker (a) systematically offers five types of therapeutic biomarker, which realized the collective consideration among different drug discovery stages, (b) comprehensively describes the therapeutic biomarker for the largest number of drugs among existing knowledge bases, and (c) extensively covers the largest classes of disease among available databases. Due to the rapid advances in AI technology, the comprehensive data collected by the marker were expected to have great implication and impact on the clinical development and clinical practice of drugs.

AIM

Identify the disease markers related to Lung cancer and categorize them into pharmacodynamic Biomarker, Monitoring biomarker, safety biomarkers

PROCEDURE

- The marker database was retrieved using following link <http://themarker.idrblab.cn/about>
- Top listed mRNA /protein Biomarker relevant to the rheumatoid arthritis data were retrieved and literature survey was done.

RESULT

B0RLD3	Cyclin-dependent kina...	Protein	CDK1; CDC2...	P	S	M	R	U	PD0166285	Lung cancer
BAJ5T6	ALK tyrosine kinase r...	Protein	ALK	P	S	M	R	U	Crizotinib	Lung cancer
B47OUL	Lactate dehydrogenase	Protein	.	P	S	M	R	U	Anti-PD1/PD-...	Lung cancer
B63BFK	Interleukin 8	Protein	CXCL8; IL8	P	S	M	R	U	Radiation ther...	Lung cancer
BQL2I7	Intercellular adhesion ...	Protein	ICAM1	P	S	M	R	U	Radiation ther...	Lung cancer
B3HTD6	C-reactive protein	Protein	CRP; PTX1	P	S	M	R	U	Radiation ther...	Lung cancer
BM7DL8	Interleukin 6	Protein	IL6; IFNB2	P	S	M	R	U	Anti-PD1/PD-...	Lung cancer
B60EZJ	Programmed cell deat...	Protein	CD274; B7H...	P	S	M	R	U	Atezolizumab	Lung cancer
B3DXG2	Serum albumin	Protein	ALB; GIG20; ...	P	S	M	R	U	Anti-PD1/PD-...	Lung cancer
B1TVG9	C-C motif chemokine 2	Protein	CCL2; MCP1...	P	S	M	R	U	Cisplatin	Lung cancer
B6CUE3	Interleukin 10	Protein	IL10	P	S	M	R	U	Anti-PD1/PD-...	Lung cancer
BUNVQ9	Interferon gamma	Protein	IFNG	P	S	M	R	U	Pembrolizumab	Lung cancer

Drug-Target Interaction Prediction: Approaches for predicting interactionsspecific to patient treatment and drug design

INTRODUCTION

Drug targets are specific molecules or biological structures involved in disease pathways that can be modulated by drugs to produce a therapeutic effect. Selecting the right drug target is crucial for the development of effective and safe medications.

Swiss Target Prediction: Swiss Target Prediction is an online tool to predict the macromolecular targets (proteins from human, mouse and rat) of bioactive small molecules. This is useful to understand the molecular mechanisms underlying a given phenotype or bioactivity, to rationalize possible side-effects, to predict off-targets and to assess the possibility of repurposing therapeutically-relevant compounds.

Swiss Target Prediction is based on the so-called similarity principle, which generally states that two similar molecules are prone to have similar properties. In the context of Swiss Target Prediction, we have statistically quantified that two similar bioactive molecules are likely to share their protein targets. Therefore, for a query molecule, we identify the most similar molecules among a set of about actives on known targets. The predicted targets are those having the actives displaying the highest similarity with the query molecule.

ABSTRACT

Repurposing drugs has become a game-changing strategy in pharmaceutical innovation, providing a tactical substitute for conventional approaches to drug development. This abstract explores the complexities of medication repurposing, including its foundational ideas, range of approaches, and significant benefits.

The understanding that already available medications, which were initially created for certain purposes, have unrealized therapeutic potential for a variety of illnesses is the cornerstone of the concept of drug repurposing. A multidisciplinary strategy incorporating knowledge from pharmacology, genetics, bioinformatics, and clinical data analysis is needed to fully realize this promise.

Several computational methods are essential to the repurposing of drugs. Large-scale omics data is analyzed using bioinformatics techniques to find possible drug-disease links, and network-based techniques reveal latent relationships between medications, targets, and illnesses.

AIM

Drug repurposing has emerged as a promising avenue to address the challenges of traditional drug discovery, offering a targeted and efficient strategy for identifying novel therapeutic uses for existing drugs. These abstract aims to delve into the intricacies of precision in drug repurposing, focusing on the specific aims and methodologies employed to achieve successful outcomes.

The primary aim of precision drug repurposing is to systematically identify and validate new therapeutic indications for existing drugs with a high degree of specificity and efficacy. This involves leveraging advanced computational techniques, including bioinformatics, network analysis, and machine learning algorithms, to prioritize potential drug-disease associations based on molecular mechanisms, genetic profiles, and clinical data.

One key aspect of precision drug repurposing is the identification of target-specific repurposing opportunities, where drugs are matched with diseases based on shared molecular pathways or targetable biomarkers. This approach allows for the rational selection of candidate drugs with the highest likelihood of success in a specific patient population or disease subtype.

It is to optimize the drug repurposing process by integrating diverse data sources and methodologies to enhance the predictive power and reliability of candidate selection. This may include the integration of multi-omics data, electronic health records, drug databases, and clinical trial data to generate comprehensive profiles of drug candidates and their potential therapeutic effects across different disease contexts.

1. INTRODUCTION

Lung cancer, also known as lung carcinoma, is a malignant tumor that begins in the lung. Lung cancer is caused by genetic damage to the DNA of cells in the airways, often caused by cigarette smoking or inhaling damaging chemicals. Damaged airway cells gain the ability to multiply unchecked, causing the growth of a tumor. Without treatment, tumors spread throughout the lung, damaging lung function. Eventually lung tumors metastasize, spreading to other parts of the body. Early lung cancer often has no symptoms and can only be detected by medical imaging. As the cancer progresses, most people experience nonspecific respiratory problems: coughing, shortness of breath, or chest pain. Other symptoms depend on the location and size of the tumor. Those suspected of having lung cancer typically undergo a series of imaging tests to determine the location and extent of any tumors. Definitive diagnosis of lung cancer requires a biopsy of the suspected tumor be examined by a pathologist under a microscope.

2. METHODOLOGY

Docking CDA with the structures 4PFJ and 6DQ4 involves computational simulations to predict their potential interactions. This process aims to determine the binding affinity and mode of interaction between CDA and each of these structures. By docking these molecules, researchers can gain insights into whether CDA could effectively target these proteins or vice versa. These insights are valuable for understanding potential therapeutic applications or guiding further research in drug development.

3. RESULT

The 2D and 3D structures illustrate the interactions between the residues of the protein and its native structure, revealing the binding affinity of these residues with the protein.

4. DISCUSSION

Genetic variations, such as the ARG20 variant in the A: ARG24 gene, can affect protein function. In the context of A: ARG24, this variant might alter the enzyme's activity, stability, or interaction with other molecules. For instance, a single amino acid change could disrupt the enzyme's catalytic site, impairing its ability to demethylate histone proteins properly. Alternatively, it could affect how the enzyme binds to its substrate or regulatory proteins, thereby influencing its function. Understanding the functional impact of the A: ARG24 variant is crucial because it provides insights into how this genetic variation might contribute to colon cancer. If the variant enhances A: ARG24 activity, it could promote the expression of genes involved in cancer progression. Conversely, if it reduces A: ARG24 activity, it might dysregulate gene expression in a way that promotes tumorigenesis.

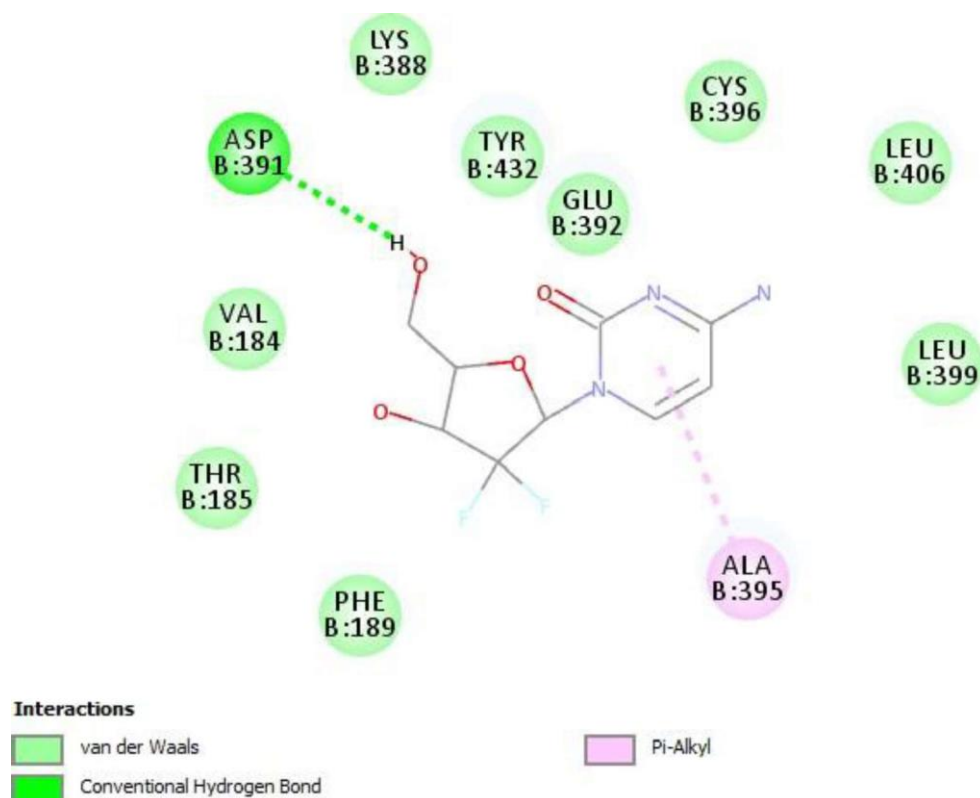
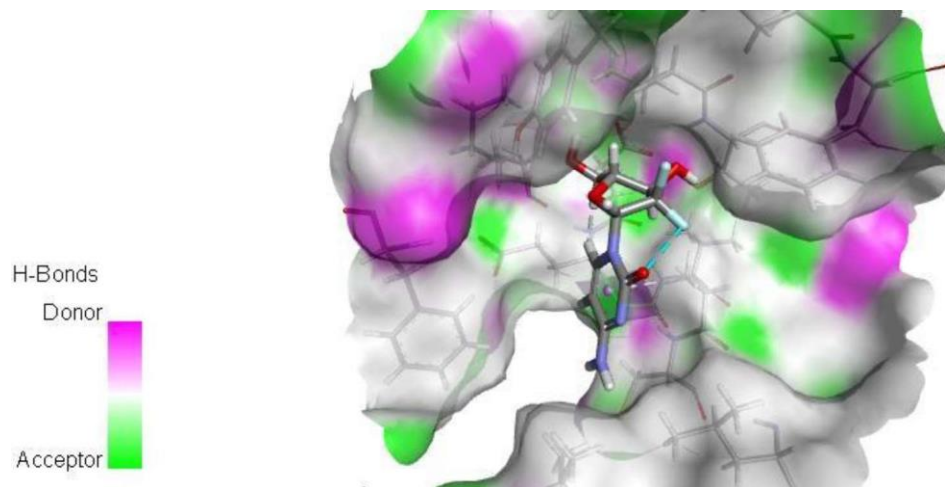
A: ARG24 has been implicated in various cancers, including colon cancer. Its role in cancer often involves regulating gene expression by demethylating histone proteins, which are crucial for controlling chromatin structure and gene transcription. In colon cancer, dysregulation of gene expression contributes to tumor development, growth, and metastasis. Research has shown that A: ARG24 expression is elevated in colon cancer tissues compared to normal colon tissues.

Moreover, increased A: ARG24 expression has been associated with poor prognosis in colon cancer patients, suggesting that it plays a role in disease progression.

5. DRUG DEVELOPMENT

Targeting A: ARG24 with drugs could be a promising approach for colon cancer therapy. By selectively inhibiting or modulating A: ARG24 activity, it might be possible to disrupt cancer-promoting pathways specifically in colon cancer cells while sparing normal cells. This targeted approach could potentially lead to more effective treatments with fewer side effects compared to traditional chemotherapeutic agents. Developing drugs that specifically target A: ARG24 requires a deep understanding of its structure, function, and regulation. Scientists can use techniques such as structural biology, high-throughput screening, and computational modelling to identify compounds that interact with A: ARG24 and modulate its activity. These compounds can then be further optimized to enhance their potency, selectivity, and pharmacokinetic properties.

2D AND 3D STRUCTURE OF 2IVX



MOLECULAR TABLE AND THEIR DOCK SCORE

GENES 1 2IVX					
MOLECULAR TABLE					
NAME	PLP1	PLP2	JAIN	PMF	DOCK SCORE
4594	25.99	27.31	-1.11	46.74	108.879
4594	21.09	27.73	-1.33	50.56	104.985
4594	8.68	14.87	-1.17	58.33	103.568
4594	4.38	8.46	-2.42	69.01	103.001
4594	4.82	6.47	-2.19	32	99.824
4594	2.78	8.1	-2.48	65.47	96.508
4594	5.51	11.82	-1.9	59.47	95.709
4594	11.35	18.89	-1.51	37.17	95.076
4594	17.91	21.72	-1.18	53.34	94.805
4594	-5.94	3.46	-1.32	58.81	93.49

NON BOND TABLE					
NAME	DISTANCE	CATEGORY	TYPES	FROM	TO
A:ARG24:NH2 - A:GLU33:OE2	2.75972	Hydrogen Bond;Electrostatic	Salt Bridge	A:ARG24:NH2	A:GLU33:OE2
A:ARG37:NH2 - A:GLU33:OE1	2.91577	Hydrogen Bond;Electrostatic	Salt Bridge	A:ARG37:NH2	A:GLU33:OE1
A:LYS92:NZ - A:GLU136:OE2	2.55916	Hydrogen Bond;Electrostatic	Salt Bridge	A:LYS92:NZ	A:GLU136:OE2
A:LYS105:NZ - A:GLU101:OE2	2.99205	Hydrogen Bond;Electrostatic	Salt Bridge	A:LYS105:NZ	A:GLU101:OE2
A:ARG258:NH2 - A:ASP168:OD1:B	3.12704	Hydrogen Bond;Electrostatic	Salt Bridge	A:ARG258:NH2	A:ASP168:OD1:B
B:ARG24:NH2 - B:GLU33:OE2	2.78089	Hydrogen Bond;Electrostatic	Salt Bridge	B:ARG24:NH2	B:GLU33:OE2
B:ARG37:NH2 - B:GLU33:OE1	2.81048	Hydrogen Bond;Electrostatic	Salt Bridge	B:ARG37:NH2	B:GLU33:OE1
B:LYS92:NZ - B:GLU136:OE2	2.71922	Hydrogen Bond;Electrostatic	Salt Bridge	B:LYS92:NZ	B:GLU136:OE2

6. MATLAB COURSES

- App Building Onramp
- Signal Processing with MATLAB
- Image Processing with MATLAB
- MATLAB Programming Techniques
- MATLAB for Data Processing and Visualization
- Machine Learning with MATLAB
- Image Processing Onramp
- Solving Nonlinear Equations with MATLAB
- Deep Learning with MATLAB
- Statistics Onramp
- MATLAB Fundamentals
- Reinforcement Learning Onramp
- Computer Vision Onramp
- Deep Learning Onramp
- Object-Oriented Programming Onramp
- Machine Learning Onramp

App Building Onramp

With its interactive, do-it-yourself design, App Building Onramp will provide quick video lessons and hands-on activities to help you learn how to use the MATLAB Application Designer to build interactive applications. After completing the application creation process, you can design, develop, and deploy MATLAB applications for a wide range of uses, from advanced computing to data visualization.

Signal Processing with MATLAB

Dive into the world of signal processing using MATLAB with the comprehensive Signal Processing Onramp. This interactive course combines theoretical concepts with hands-on exercises to help you understand and apply various signal processing techniques using MATLAB. From filtering and spectrum analysis to signal imaging, this Onramp gives you the capabilities to analyse and manipulate signals for a variety of applications.

Image Processing with MATLAB

Unleash the power of MATLAB for image manipulation with our Onramp image editor. This hands-on course is designed to introduce you to the basics of image processing and show you how MATLAB can be used to perform a variety of image analysis tasks. Through interactive tutorials and real-world examples, you will learn techniques such as image enhancement, segmentation and feature extraction that enable you to efficiently process and interpret images.

MATLAB Programming Techniques

Take your MATLAB skills to the next level with MATLAB Programming Techniques Onramp. This advanced course is designed for experienced MATLAB users who want to improve their programming skills. From optimization and debugging strategies to designing advanced data structures and algorithms, this Onramp covers a wide range of programming techniques to help you write powerful, robust, and maintainable MATLAB.

MATLAB for Data Processing and Visualization

It's a good idea to learn MATLAB for data processing and visualization, especially if you work with numerical data. Here are some things you might want to learn about processing and visualizing a data with MATLAB. These MATLAB capabilities enable you to handle a variety of data processing and visualization tasks, making it a powerful tool for data analysis and visualization.

Machine Learning with MATLAB

MATLAB is a versatile machine learning environment that contains rich built-in tools, functions, and documentation. Gaining knowledge of MATLAB machine learning will enable you to apply various machine learning techniques to real-world problems.

Image Processing Onramp

With an interactive, do-it-yourself approach, Image Processing Onramp offers short video courses and hands-on exercises to help you understand and use MATLAB image processing concepts. Complete the image processing onramp for a thorough understanding of image processing principles and methods. You will also gain hands-on experience using MATLAB to develop image processing algorithms.

Solving Nonlinear Equations with MATLAB

Unleash the power of MATLAB to solve complex nonlinear equations with Nonlinear Equation Solving Onramp. This self-paced course is designed to give you the skills you need to effectively handle nonlinear equations using the extensive computing capabilities of MATLAB. Through guided exercises, you will learn various numerical methods and techniques for finding solutions to nonlinear equations in various scientific and engineering applications.

Deep Learning with MATLAB

Explore the advancement of deep learning with MATLAB with Deep Learning Onramp. This interactive course provides a thorough introduction to deep learning concepts and shows how to use MATLAB to build, train, and deploy deep neural networks. Whether you're new to deep learning or want to deepen your knowledge, this Onramp gives you the skills to tackle complex problems.

Statistics Onramp

Dive into the world of statistical analysis and data interpretation with MATLAB Statistics Onramp. This interactive learning experience provides a comprehensive overview of statistical concepts, from basic descriptive statistics to advanced data modelling techniques. By learning the statistical capabilities of MATLAB, you can analyse data, generate meaningful insights, and make informed decisions in a variety of fields.

MATLAB Fundamentals

Learn the basics of MATLAB programming from the ground up with MATLAB Fundamentals Onramp. Designed for beginners, this interactive course provides a structured learning path covering basic MATLAB syntax, data types, control flow, and function writing. Whether you're new to programming or transitioning from another language, this Onramp provides a solid foundation for your MATLAB journey.

Reinforcement Learning Onramp

Reinforcement Learning Onramp for MATLAB is a comprehensive introductory course that teaches you how to use MATLAB to build reinforcement learning algorithms and the principles of reinforcement learning.

Computer Vision Onramp

Computer vision is where knowledge from the real world meets digital data. Data from multiple cameras and sensors can be combined to create a digital reconstruction. In this course, you will use object recognition to find, track, and count turtles in a video.

Deep Learning Onramp

Deep Learning Onramp is an interactive, self-paced course that uses MATLAB to teach you how to apply and understand deep learning concepts through short video tutorials and hands-on activities. After completing the Deep Learning Onramp, you will have a thorough understanding of deep learning theories and methods and hands-on experience building and improving deep learning models in MATLAB.

Object-Oriented Programming Onramp

Journey through object-oriented programming (OOP) principles with the MATLAB object-oriented programming onramp. This interactive course combines concise video tutorials with hands-on exercises to teach you the basics of OOP in MATLAB. By the end of this Onramp, you will have a good understanding of how to use OOP concepts to create modular and reusable code in MATLAB.

Machine Learning Onramp

Journey through object-oriented programming (OOP) principles with the MATLAB object-oriented programming onramp. This interactive course combines concise video tutorials with hands-on exercises to teach you the basics of OOP in MATLAB. By the end of this Onramp, you will have a good understanding of how to use OOP concepts to create modular and reusable code in MATLAB.



Course Completion Certificate

Praveenraja K

has successfully completed **100%** of the self-paced training course

Signal Processing with MATLAB



DIRECTOR, TRAINING SERVICES

6 April 2024



Course Completion Certificate

Praveenraja K

has successfully completed **100%** of the self-paced training course

Deep Learning with MATLAB



DIRECTOR, TRAINING SERVICES

9 February 2024



Course Completion Certificate

Praveenraja K

has successfully completed **100%** of the self-paced training course

Deep Learning with MATLAB



Craig L. Santos

DIRECTOR, TRAINING SERVICES

9 February 2024



Course Completion Certificate

Praveenraja K

has successfully completed **100%** of the self-paced training course

MATLAB Programming Techniques



Craig L. Santos

DIRECTOR, TRAINING SERVICES

17 February 2024



Course Completion Certificate

Praveenraja K

has successfully completed **100%** of the self-paced training course

Machine Learning with MATLAB


DIRECTOR, TRAINING SERVICES

12 February 2024



Course Completion Certificate

Praveenraja K

has successfully completed **100%** of the self-paced training course

Simulink Onramp


DIRECTOR, TRAINING SERVICES

28 March 2024



Course Completion Certificate

Praveenraja K

has successfully completed **100%** of the self-paced training course

Image Processing with MATLAB



DIRECTOR, TRAINING SERVICES

27 March 2024



Course Completion Certificate

Praveenraja K

has successfully completed **100%** of the self-paced training course

Computer Vision Onramp



DIRECTOR, TRAINING SERVICES

21 March 2024



Course Completion Certificate

Praveenraja K

has successfully completed **100%** of the self-paced training course

Solving Nonlinear Equations with MATLAB


DIRECTOR, TRAINING SERVICES

21 March 2024



Course Completion Certificate

Praveenraja K

has successfully completed **100%** of the self-paced training course

MATLAB for Data Processing and Visualization


DIRECTOR, TRAINING SERVICES

14 March 2024



Course Completion Certificate

Praveenraja K

has successfully completed **100%** of the self-paced training course

Object-Oriented Programming Onramp


DIRECTOR, TRAINING SERVICES

5 March 2024

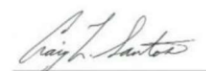


Course Completion Certificate

Praveenraja K

has successfully completed **100%** of the self-paced training course

Machine Learning Onramp


DIRECTOR, TRAINING SERVICES

29 February 2024

Unravelling Biological Complexity Using Machine

Abstract:

This article examines the developing field of using machine learning methods to decipher biological systems' intricate workings. Technological developments are leaps and bounds of biological data, yet conventional methods of analysis are not able to adequately capture the complexities of biological processes. Deeper understanding of biological processes is made possible by machine learning algorithms, which provide a potential method for removing meaningless patterns and insights from these large, complicated datasets.

Introduction:

The intricacy of life processes is beyond our current comprehension. Deciphering the complexities of these systems, which range from the dynamics of ecosystems to the molecular interactions within cells, is a formidable task. Conventional reductionist methods have shed important light on particular facets of biological phenomena. Nevertheless, they frequently fall short of capturing the interdependence of their constituent parts and the holistic character of biological systems.

Application of Machine Learning in Systems Biology:

With the availability of more data and increased processing capacity, machine learning approaches have become increasingly useful for handling the complexity of biological systems. These methods make use of algorithms that don't require explicit programming in order to learn from data, spot trends, and generate predictions. Machine learning techniques are being used in systems biology in a number of different areas, such as systems ecology, proteomics, metabolomics, and genomics.

Genomic Analysis:

Large-scale genomic datasets are analysed in genomics using machine learning methods, which can identify patterns in genomic variation, regulatory networks, and gene expression. These algorithms are able to determine the genetic basis of complex features, forecast gene function, and uncover biomarkers linked to disease.

Proteomic and Metabolomic Studies:

Similar to this, machine learning methods are used in proteomics and metabolomics to interpret the intricate connections between proteins, metabolites, and biological pathways. These techniques allow the identification of important molecular actors participating in different physiological processes and disease states by integrating multi-omics data.

Systems Ecology and Environmental Biology:

Machine learning techniques are used in systems ecology and environmental biology to model and forecast species interactions, ecosystem dynamics, and environmental reactions to perturbations. These models support conservation efforts and provide information for policy decisions by illuminating the intricate interactions that exist between creatures and their habitats.

Challenges and Future Directions:

Even with the enormous advancements, there are still obstacles in the way of fully utilising machine learning to decipher biological complexity. Significant obstacles include things like data quality, model interpretability, and ethical considerations. It will take interdisciplinary teamwork, strict validation procedures, and moral standards for the appropriate application of machine learning in biology to overcome these obstacles.

Conclusion:

To sum up, the amalgamation of machine learning methodologies with systems biology presents unparalleled prospects for deciphering the intricacies of biological systems. Through the utilisation of data-driven methodologies, scientists can acquire a more profound understanding of basic biological processes, resulting in revolutionary breakthroughs in the fields of environmental sustainability, agriculture, and healthcare.

Praveen raja

by OfficialTurnitin

Submission date: 30-Apr-2024 10:03PM (UTC+0900)

Submission ID: 2366697424

File name: Untitled-document-21.docx (287.2K)

Word count: 484

Character count: 3221

Unravelling Biological Complexity Using Machine Learning

Abstract:

This article examines the developing field of using machine learning methods to decipher biological systems' intricate workings. Technological developments are making it possible to generate enormous amounts of biological data, yet conventional methods of analysis are not able to adequately capture the complexities of biological processes. Deeper understanding of biological processes is made possible by machine learning algorithms, which provide a potential method for removing meaningless patterns and insights from these large, complicated datasets.

Introduction:

The intricacy of life processes is beyond our current comprehension. Deciphering the complexities of these systems, which range from the dynamics of ecosystems to the molecular interactions within cells, is a formidable task. Conventional reductionist methods have shed important light on particular facets of biological phenomena. Nevertheless, they frequently fall short of capturing the interdependence of their constituent parts and the holistic character of biological systems.

Application of Machine Learning in Systems Biology:

With the availability of more data and increased processing capacity, machine learning approaches have become increasingly useful for handling the complexity of biological systems. These methods make use of algorithms that don't require explicit programming in order to learn from data, spot trends, and generate predictions. Machine learning techniques are being used in systems biology in a number of different areas, such as systems ecology, proteomics, metabolomics, and genomics.

Genomic Analysis:

Large-scale genomic datasets are analysed in genomics using machine learning methods, which can identify patterns in genomic variation, regulatory networks, and gene expression. These algorithms are able to determine the genetic basis of complex features, forecast gene function, and uncover biomarkers linked to disease.

Proteomic and Metabolomic Studies:

Similar to this, machine learning methods are used in proteomics and metabolomics to interpret the intricate connections between proteins, metabolites, and biological pathways. These techniques allow the identification of important molecular actors participating in different physiological processes and disease states by integrating multi-omics data.

Systems Ecology and Environmental Biology:

Machine learning techniques are used in systems ecology and environmental biology to model and forecast species interactions, ecosystem dynamics, and environmental reactions to perturbations. These models support conservation efforts and provide information for policy decisions by illuminating the intricate interactions that exist between creatures and their habitats.

Challenges and Future Directions:

Even with the enormous advancements, there are still obstacles in the way of fully utilising machine learning to decipher biological complexity. Significant obstacles include things like data quality, model interpretability, and ethical considerations. It will take interdisciplinary teamwork, strict validation procedures, and moral standards for the appropriate application of machine learning in biology to overcome these obstacles.

Conclusion:

To sum up, the amalgamation of machine learning methodologies with systems biology presents unparalleled prospects for deciphering the intricacies of biological systems. Through the utilisation of data-driven methodologies, scientists can acquire a more profound understanding of basic biological processes, resulting in revolutionary breakthroughs in the fields of environmental sustainability, agriculture, and healthcare.

Praveen raja

ORIGINALITY REPORT

3%

SIMILARITY INDEX

0%

INTERNET SOURCES

0%

PUBLICATIONS

3%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to Berlin School of Business and Innovation

Student Paper

3%

Exclude quotes Off

Exclude bibliography On

Exclude matches Off

Praveen raja

GRADEMARK REPORT

FINAL GRADE

GENERAL COMMENTS

/0

PAGE 1

PAGE 2