

INTEGRATED ANALYSIS: GENE EXPRESSION, SURVIVAL, BIOMARKERS AND DRUG TARGETS

Submitted in partial fulfilment of the required for the award of

**Master of Science Degree in Bioinformatics
and Data Science by SHYAM A (43739030)**



**DEPARTMENT OF
BIOINFORMATICS AND DATA SCIENCE**

**SCHOOL OF BIO AND CHEMICAL
ENGINEERING**

SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)

**Accredited with Grade “A++” by NAAC JEPPIAAR
NAGAR, RAJIV GANDHI SALAI, CHENNAI – 600119**

APRIL 2024



SATHYABAM_A

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)

Accredited with “A” grade by NAAC I 12B Status by UGC I Approved by AICTE

Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai – 600 119

www.sathyabama.ac.in

DEPARTMENT OF BIOINFORMATICS AND DATA SCIENCE

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the Bonafide work of SHYAM A (43739030) who carried out the project entitled “Integrated Analysis: Gene Expression, Survival, Biomarkers, and Drug Targets” under my / our supervision from January 2024 to April 2024

INTERNAL GUIDE

External Guide (if Applicable)

Head of the Department

Submitted for Viva-voce Examination held on _____

INTERNAL EXAMINER

EXTERNAL EXAMINER

DECLARATION

I, **SHYAM A** hereby declare that the project entitled “Integrated Analysis: Gene Expression, Survival, Biomarkers, and Drug Targets” done by me under the guidance of **DR.D.ALEX ANAND** is submitted in partial fulfilment of the requirements for the award of Master of Science degree in Bioinformatics and Data Science.

DATE:

SHYAM. A

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 CHAKRAVARTHI M.E., Ph.D. Dean of Arts and Science** for providing me necessary support and details at the right time during the progressive reviews.

I would like to express my sincere and deep sense of gratitude to my project guide

DR. D. ALEX ANAND Department of Bioinformatics and Data Science for his valuable guidance, suggestions, and constant encouragement that paved way for the successful completion of my project work.

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 for the completion of the project.

(SHYAM)

LITERATURE REVIEW

Title: "Harnessing Machine Learning in Systems Biology"

Abstract:

Systems biology, a subject dedicated to thoroughly understanding biological systems, has seen a significant increase in data output as a result of developments in high-throughput technologies. Nevertheless, the size and complexity of these datasets have made it extremely difficult to draw conclusions that are insightful. Machine learning algorithms are becoming very effective tools for processing this kind of data, providing new ways to find patterns, forecast results, and clarify biological principles. This paper examines how machine learning is incorporated into systems biology, emphasising how it has revolutionised biological research and outlining both present issues and potential future directions.

Introduction:

High-throughput technologies have completely changed biological research by making it possible to generate enormous amounts of data on a wide range of biological scales, from molecular interactions to the behaviors of entire organisms. Nevertheless, the intricacy and variability of these datasets frequently prove to be too much for conventional analytical techniques to handle, which hinders our capacity to glean significant biological insights. The goal of the science of systems biology, which evolved in response to this problem, is to comprehend biological systems as integrated networks of interacting components. Integrating experimental data with computational models is essential to systems biology because it gives researchers a comprehensive understanding of biological processes. Nevertheless, the sheer amount and complexity of biological data provide a significant analytical problem, calling for the creation of novel computing strategies. Within this framework, machine learning has become

Machine Learning in Systems Biology:

¹ Without explicit programming, computers can learn from data and make predictions or judgments thanks to a variety of computational processes known as machine learning. Machine learning techniques can be used to analyze several kinds of biological data in the setting of systems biology, including transcriptomics, proteomics, metabolomics, imaging, genomics, and others. By removing patterns and relationships from large, complicated datasets, these algorithms help scientists learn new things about biological processes. In the field of genomics, machine

learning algorithms have been employed to detect genetic variants linked to various diseases, forecast the roles of genes, and categorize distinct cell types according to their gene expression patterns. Similar to this, machine learning methods have been used in structural biology to forecast protein structures, deduce protein-protein interactions, and create new therapies.

Applications and Impact:

Applications in many different fields of biological research have resulted from the integration of machine learning into systems biology. Machine learning algorithms are used in drug development and discovery to select drug targets, create new treatments with increased potency and specificity, and forecast the safety and efficacy of candidate molecules. Machine learning algorithms examine patient data in customized medicine to forecast the likelihood of illness, customize treatment plans, and improve patient outcomes. Machine learning techniques are used in agriculture to improve crop breeding programs, forecast disease outbreaks, and maximize crop output and quality. In ecology, species distributions are modeled, ecosystem dynamics are predicted, and conservation policies are informed by machine learning algorithms that evaluate environmental data.

Machine learning is having an impact on systems biology that goes beyond disciplinary boundaries and promotes interdisciplinary

Future Directions:

In terms of expanding our knowledge of biological systems and tackling urgent issues in agriculture, the environment, and health, machine learning in systems biology has a bright future. Our capacity to model and simulate complex biological systems with previously unheard-of accuracy and granularity will be further improved by ongoing developments in machine learning algorithms in conjunction with advances in data collecting and integration methods. Furthermore, novel directions for investigating biological complexity at the cellular and molecular levels will become possible with the combination of machine learning with cutting-edge technologies like spatial transcriptomics, single-cell sequencing, and multi-omics methodologies.

Conclusion:

In summary, a paradigm change in our knowledge of biological systems is brought about by the incorporation of machine learning into systems biology. Researchers can find hidden patterns, forecast biological events, and hasten scientific discovery by using machine learning algorithms to examine massive biological datasets. But there are a number of methodological, technological, and ethical obstacles that must be overcome before machine learning in systems biology can reach its full potential. However, the complementary fields of systems biology and machine learning have great potential to transform biological research and tackle some of the most important problems confronting humanity in the twenty-first century.

Shyam A

ORIGINALITY REPORT

4%

SIMILARITY INDEX

1%

INTERNET SOURCES

0%

PUBLICATIONS

3%

STUDENT PAPERS

PRIMARY SOURCES

1

Submitted to Heriot-Watt University

Student Paper

3%

2

www.mdpi.com

Internet Source

1%

Exclude quotes Off

Exclude matches Off

Exclude bibliography On

MINIPROJECT

INTRODUCTION

A Biological network represents systems as complex set of binary interactions or relation between various biological systems and can reveal informative patterns within the systems. Cytoscape is a open source software platforms for visualizing molecule interactions network and

biological pathway and integrating the network with annotations, gene expression profile. Cytohubba predict and explore the important nodes and edges in given network by several topological algorithm such as degree, bottleneck, MCC, DMNC, MNC, EPCv, Eccentricity, Closeness, Stress, Clustering co-efficient and betweenness.

ABSTRACT

This study aimed to elucidate the molecular mechanisms underlying biological processes or disease conditions through the analysis of gene expression data retrieved from the Gene Expression Omnibus (GEO) database and visualization of the results in 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 AND METHODS

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's 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

CytoscapeS 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's an overview of CytoHubba and its features.

CLUEGO

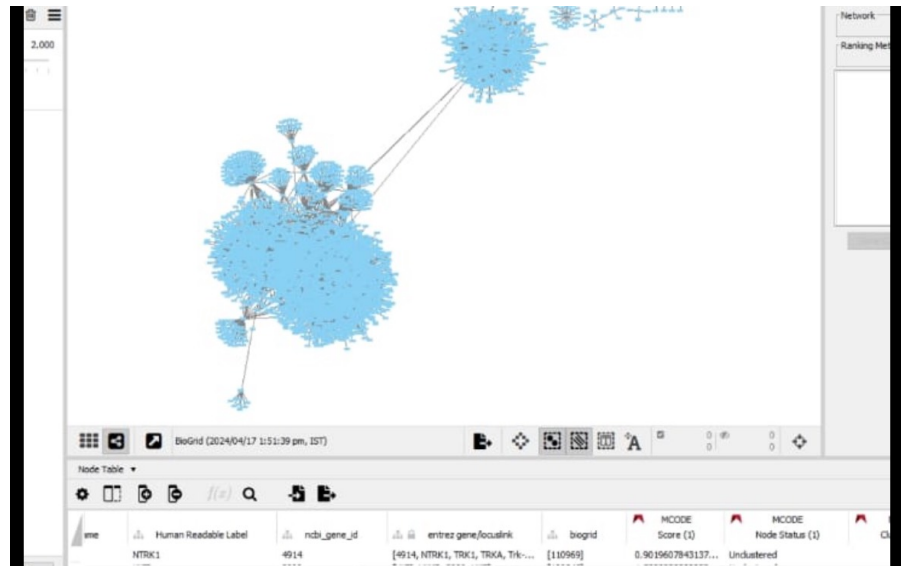
ClueGO is a Cytoscape plugin designed to analyze and visualize functional enrichment of gene ontology (GO) terms and biological pathways within biological networks. It is commonly used by researchers in bioinformatics and systems biology to gain insights into the biological significance of gene sets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

PROCEDURE

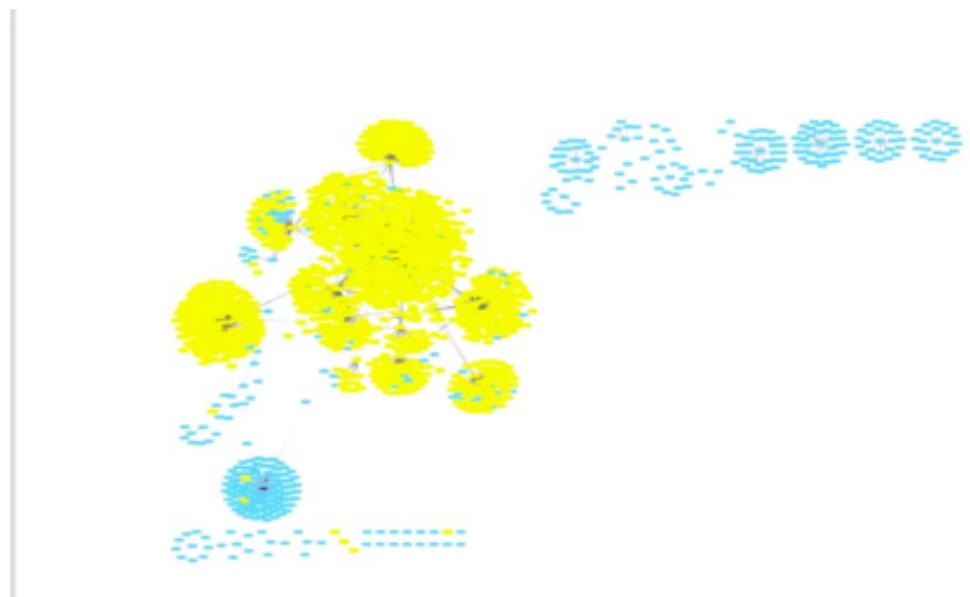
- For the database and disease pulmonary fibrosis Geo2R is the source of the geo dataset (GSE GSE81558)
- After choosing the suitable parameters, the Geo dataset is retrieved.
- Genes that were upregulated genes were copied and paste into the cytoscapae network.
- Built a biological network using a BioGrid in cytoghubba to analyse the network and identify importance node based various topologies algorithm to determine the the top 10 gens(IBMPFD3,ALS21,A1U,ALPS2B,Actr1a,UBC,TCF-3C,A1,PHA2,ENTRK1)
- Genes must be copied in cluego and contigue cluego parameter such as species ontology and statistical settings.
- Run cluego analysis to identify functional categories or pathways associated with the selected nodes.

RESULT

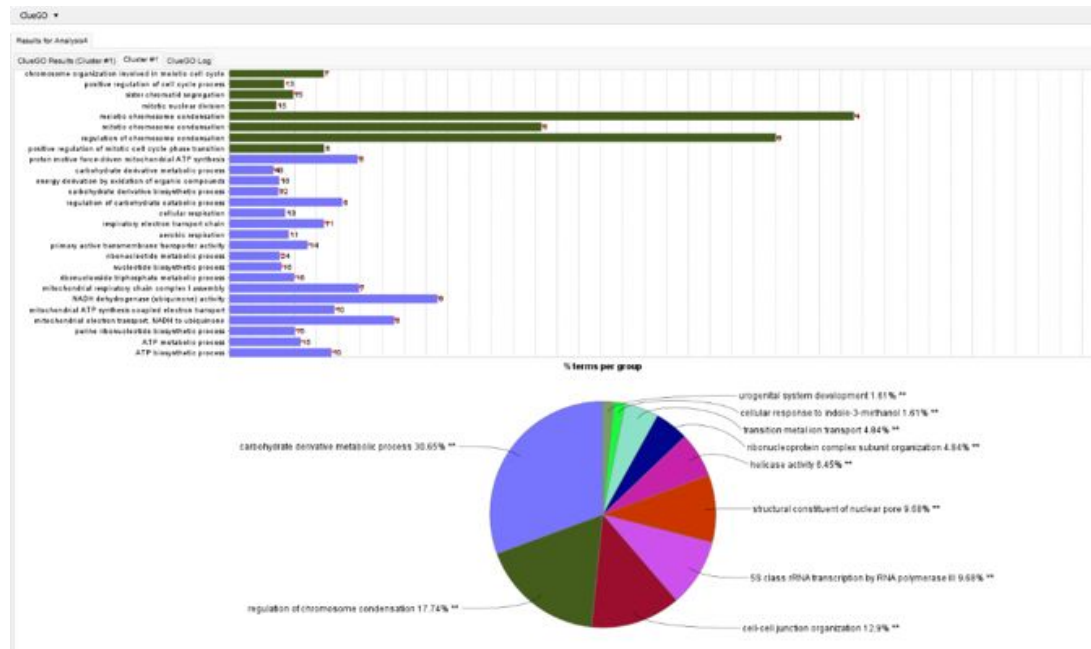
Cytoscape



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 approach for unraveling the molecular underpinnings of biological processes and disease conditions. By leveraging publicly available data and network-based visualization tools, researchers can gain actionable insights into gene regulation, pathway dynamics, and disease mechanisms, driving forward our understanding of complex biological systems.

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 PRB1 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 0.79(0.97-2.52) 50% median survival between groups . it represent the ratio of probabilityfor low expression cohort (25.93)and high expression cohort(55.67)

BREAST CANCER

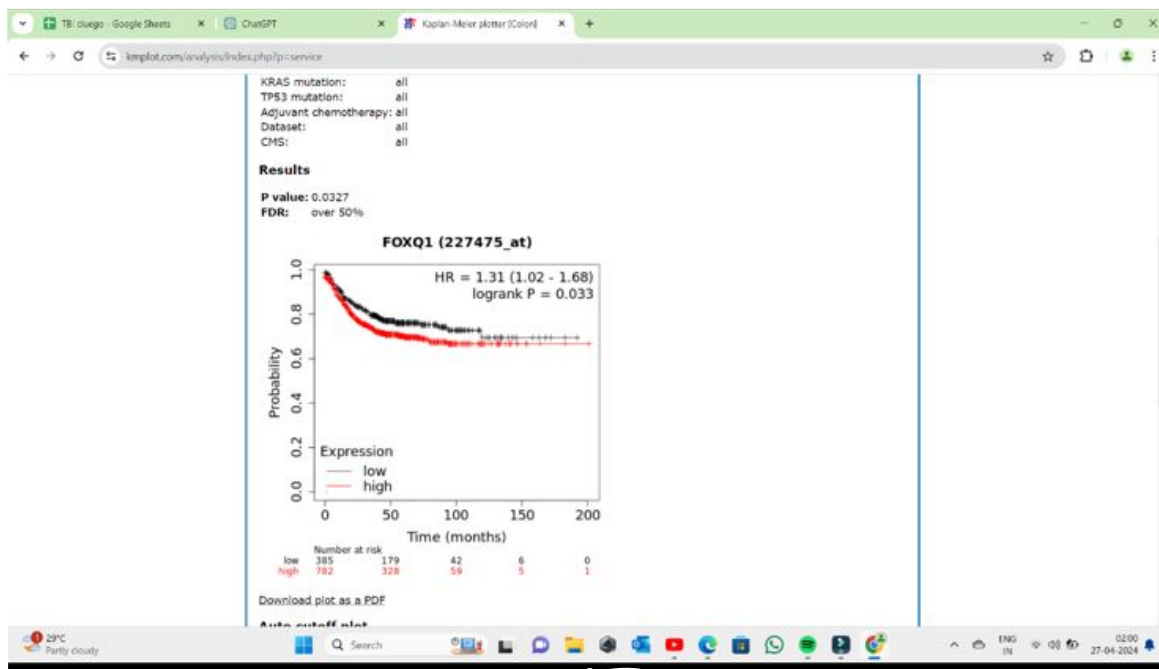
The HR of MX1 in breast cancer is 2.58(1.84-3.61)50% median survival between groups . it represent the ratio of probability for low expression cohort (215.2) and high expression cohort(93.27)

CERVICAL SQUAMOUS CELL CARCINOMA

The HR of MXI in cervical squamous cell carcinoma is 1.57(0.97-2.52)50% median survival between groups . it represent the ratio of probability for low expression cohort (37.27)and high expression cohort(20.23)

ESOPHAGAL SQUAMOUS CELL

The HR of MXI in esophagal squamous cell is 1.95(0.82-4.6)50% median survival between groups . it represent the ratio of probability for low expression cohort (45.37)and high expression cohort(25.43)



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's were retrieved and literature survey was done.

The screenshot shows a web browser displaying the OncoPrint database search results for the query 'Leukemia biomarker'. The results are presented in a table with columns for Biomarker ID, Biomarker name, class, Gene name, Biomarker type, Representative drug, and Representative disease. The table lists various biomarkers such as BMH7C, BNOP10, B1ZX4, B7V0QX, BVP58M, B9U7QX, BMH01Z, BW8T6O, B25K3P, BFK3V9, BTPK19, B57IMO, BT9QP3, BLC2F3, and BZ32NV, each associated with a specific gene and a representative drug.

Biomarker ID	Biomarker name	class	Gene name	Biomarker type	Representative drug	Representative disease
BMH7C	Thiopurine S-methyltr...	Protein	TPMT	P S M R U	Thioguanine	Acute myeloid l...
BNOP10	Cystatin C	Protein	CST3	P S M R U	Furosemide	Healthy
B1ZX4	T-lymphocyte activatio...	Protein	CD86; CD26...	P S M R U	Fluorinated tra...	Acute myeloid l...
B7V0QX	Tyrosine-protein kinas...	Protein	ABL1; ABL; J...	P S M R U	Nilotinib	Parkinsonism
BVP58M	Crk-like protein	Protein	CRKL	P S M R U	Imatinib	Chronic myeloi...
B9U7QX	Leukocyte surface ant...	Protein	MS4A1; CD20	P S M R U	Rituximab	Acute lymphobl...
BMH01Z	Myeloid cell surface a...	Protein	CD33; SIGL...	P S M R U	Gemtuzumab ...	Acute myeloid l...
BW8T6O	B-lymphocyte surface ...	Protein	CD19	P S M R U	Blinatumomab	Acute lymphobl...
B25K3P	Ferritin	Protein	FTMT	P S M R U	Defeniprone	Parkinsonism
BFK3V9	CNT3 mRNA	mRNA	SLC28A3; C...	P S M R U	Cetuximab	Head and neck...
BTPK19	HEXIM1 mRNA	mRNA	HEXIM1; CL...	P S M R U	BI 894999	Acute myeloid l...
B57IMO	HMGCLL1	Gene	DKFZP434G...	P S M R U	Imatinib	Chronic myeloi...
BT9QP3	UGT1A1	Gene	UGT1A	P S M R U	Nilotinib	Chronic myeloi...
BLC2F3	IDH2	Gene	IDH-2	P S M R U	Enasidenib	Acute myeloid l...
BZ32NV	IDH1	Gene		P S M R U	Ivosidenib	Liver cancer

INTERPRETATION:

Biomarker ID	Biomarker name	class	G e n e name	Biomarker type	Representative drug	Representative disease
BPV52Q	T u m o r necrosis factor	protein	T N F ; TNFA; T	PSMR	Infliximab	Rheumatoid arthritis

Drug-Target Interaction Prediction: Approaches for predicting interactions specific 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 quantify 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 370'000 actives on known targets. The predicted targets are those having the actives displaying the highest similarity with the query molecule.

AIM:

Drug Benzalkonium targets screening and binding assessment.

MATERIALS AND METHODS:

Swiss Target Prediction:

SwissTargetPrediction is an online tool to predict the targets of bioactive small molecules in human and other vertebrates. This is useful to understand the molecular mechanisms underlying a given phenotype or bioactivity, to rationalize possible side-effects or to predict off-targets of known molecules.

Uniprot: UniProt is a freely accessible database of protein sequence and functional information, many entries being derived from genome sequencing projects. It contains a large amount of information about the biological function of proteins derived from the research literature. It is maintained by the UniProt consortium, which consists of several European bioinformatics organisations and a foundation from Washington, DC, USA.

PDB: The Protein Databank (PDB) is a file format describing the 3D structure of molecules held in protein data bank. The PDB format accordingly provides for description and annotations of protein and nucleic acid structures including atomic co-ordinates, secondary structure assignments as well as atomic connectivity.

Drug Bank: Drug bank database is a comprehensive, freely accessible online database containing information on drugs and drug targets. This database is a unique bioinformatics and cheminformatics resource that combines detailed drug.

RESULT

MOLECULE TABLE:

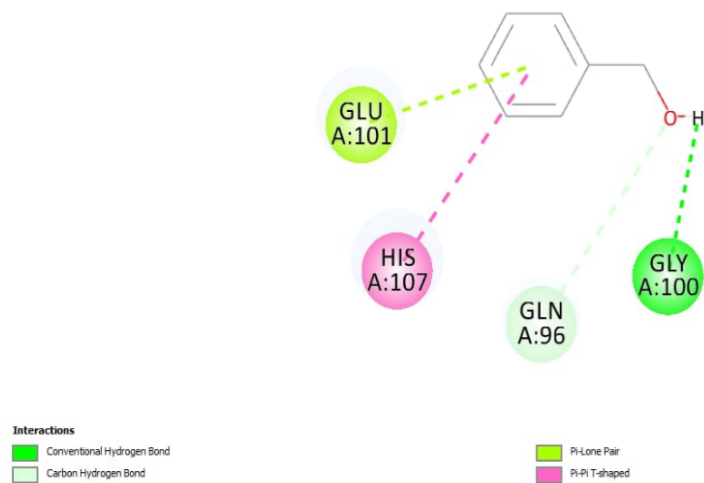
						DO CK _S CO RE
NA ME	-PL P1	-PL P2	JAI N	-P MF		
	31. 14	36. 9	-0. 19	5.1 2		20. 226
244	31. 14	36. 9	-0. 19	5.1 2		20. 226
244	31. 14	36. 9	-0. 19	5.1 2		20. 226
	31. 79	37. 21	0.1 7	6.1 4		20. 22
244	32. 19	37. 81	-0. 15	5.6 8		20. 102
244	36. 39	41. 43	-0. 17	6.8 5		20. 028
244	36. 39	41. 43	-0. 17	6.8 5		20. 028
244	36. 16	41. 54	0.0 7	5.9 8		20. 025
244	36. 16	41. 54	0.0 7	5.9 8		20. 025
244	36. 59	42. 01	0.0 2	7.6 7		20. 012

NON-BONDING TABLE:

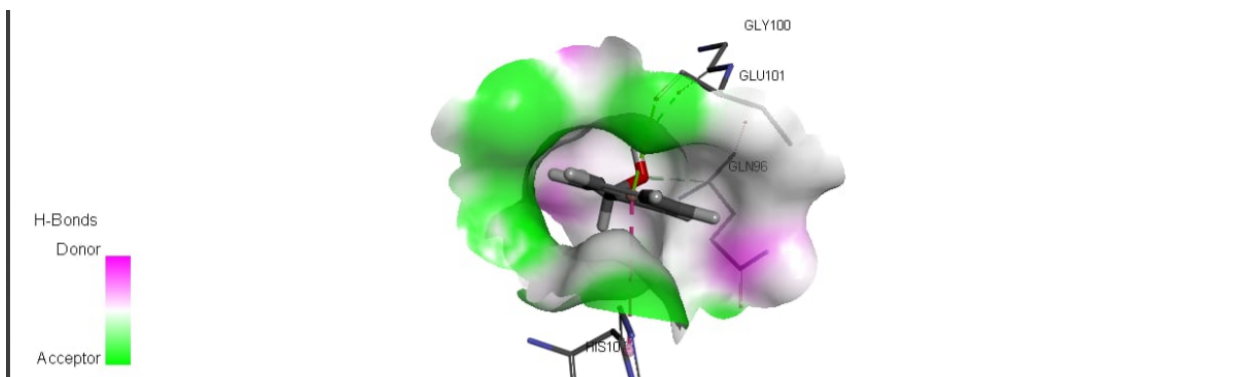
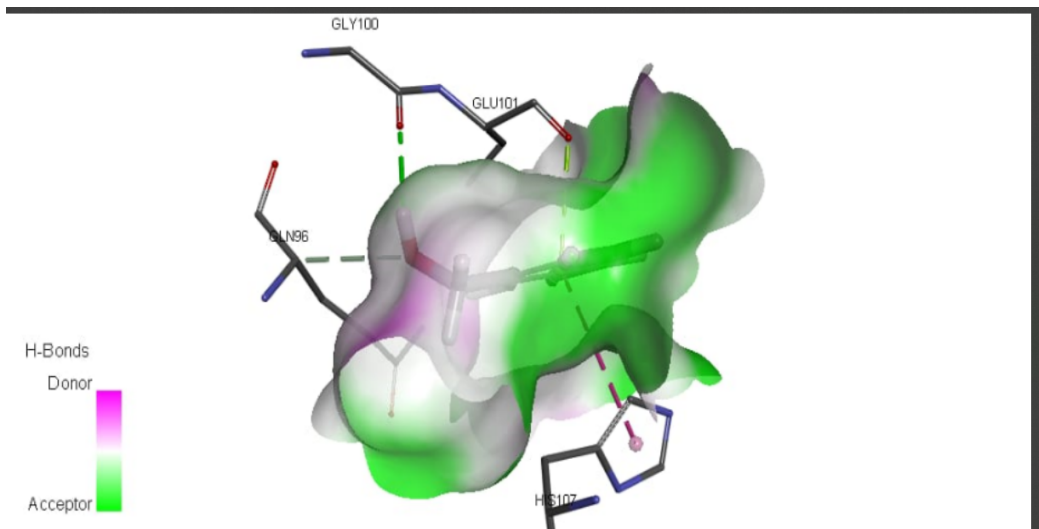
	DI ST AN CE	CA TE GO RY		TY PE	FR O M	TO
NA ME						

244 :H1 6 - A: GL Y1 00: O	2.2 584 8	Hy dro gen Bo nd	Co nve nti ona l Hy dro gen Bo nd	244 :H1 6	A: GL Y1 00: O
A: GL N9 6:C A - 244 :O1	3.2 446 3	Hy dro gen Bo nd	Car bon Hy dro gen Bo nd	A: GL N9 6:C A	244 :O1
A: GL U1 01: O - 244	2.8 928 9	Oth er	Pi- Lo ne Pai r	A: GL U1 01: O	244
A: HI S1 07 - 244	4.5 415 2	Hy dro pho bic	Pi- Pi T- sha ped	A: HI S1 07	244

2D STRUCTURE:



3D STRUCTURE:



INTERPRETATION

Benzalkonium chloride (BZK) is commonly used in pharmaceuticals as an antimicrobial agent. It's found in various over-the-counter medications, including throat lozenges, nasal sprays, and eye drops. In these formulations, benzalkonium chloride helps prevent the growth of bacteria and other microorganisms, which can contaminate the product or cause infections. JM21 refers to a gene, it would be helpful to know its function or relevance to understand any potential relationship with benzalkonium chloride.

Course Certificates

Based on the matlab Certificate there are 17 Course I have completed

MATLAB Onramp

MATLAB for Data Processing and Visualization

Machine Learning with MATLAB

Computer Vision Onramp

Reinforcement Learning Onramp

Image Processing with MATLAB

Wireless Communications Onramp

Deep Learning Onramp

Signal Processing with MATLAB

Solving Nonlinear Equations with MATLAB

MATLAB Programming Techniques

Statistics Onramp

Image Processing Onramp

Object-Oriented Programming Onramp

Machine Learning Onramp

MATLAB Fundamentals

Deep Learning with MATLAB

MATLAB ONRAMP

INTRODUCTION:

MATLAB Onramp is a free, interactive tutorial that helps beginners learn MATLAB, a high-level programming language and environment used in technical computing and data visualization. It covers the basics of MATLAB syntax, programming, and problem-solving techniques. It's a great starting point for anyone new to MATLAB.

WHAT WE LEARNT

1. Introduction to Matlab
2. Basic operations
3. Vectors and matrices
4. Plotting and visualisation
5. Programming concepts
6. Advanced data types
7. File I/O
8. Symbolic mathematics
9. Applications

DURATION: 2hrs

GRADE: 100%

2. OBJECT ORIENTATED PROGRAMMING ONRAMP

INTRODUCTION:

"Object-Oriented Programming Onramp" likely refers to a learning path or course designed to introduce individuals to the fundamentals of object-oriented programming (OOP). In OOP, programs are organized around objects, which are instances of classes that encapsulate data and behavior. It's a popular programming paradigm used in many languages like Java, C++, and Python.

WHAT WE LEARNT

1. Introduction of OOP
2. Classes and object
3. Inheritance
4. Polymorphism
5. Encapsulation
6. Abstraction
7. Method overriding and overloading
8. Interfaces and abstract classes
9. Design principles
- 10 Practical examples and exercises

DURATION: 2hrs

GRADE: 100%

3. MATLAB FUNDAMENTALS

INTRODUCTION:

MATLAB Fundamentals typically refer to a course or material that covers the basics of MATLAB, a high-level programming language and environment primarily used for numerical computing and data analysis. These fundamentals may include topics like syntax, data types, matrices, plotting, functions, and basic programming concepts tailored to MATLAB's unique features and applications.

WHAT WE LEARNT

1. Basic Syntax and commands
2. Variable and data types
3. Basic operations and functions
4. Control flow statements
5. Array manipulation and indexing
6. Plotting and visualisation
7. Basic debugging techniques

DURATION: 16.5 hrs

GRADE: 100%

4. MACHINE LEARNING ONRAMP

INTRODUCTION:

"MACHINE LEARNING ONRAMP" likely refers to a structured learning path or course designed to provide an introduction to machine learning concepts and techniques. This onramp could cover foundational topics such as supervised learning, unsupervised learning, regression, classification, clustering, and evaluation methods. It might also include hands-on exercises or projects to help learners gain practical experience in applying machine learning algorithms to real-world problems.

WHAT WE LEARNT

1. Introduction machine learning concepts and terminology
2. Data preprocessing techniques such as cleaning normalisation and features scaling
3. Supervised learning algorithms including linear regression, logistic regression, decision tree and k nearest neighbour
4. Evaluation metrics for assessing model performance
5. Over fitting, under fitting and techniques for model validation and selection

DURATION: 2hrs

GRADE: 100%

5. MACHINE LEARNING WITH MATLAB

INTRODUCTION

"Machine Learning with MATLAB" is a comprehensive approach to mastering the principles and applications of machine learning using MATLAB, a powerful programming language and environment. In this introduction, you'll explore how MATLAB provides a versatile platform for implementing various machine learning algorithms, from classical techniques to cutting-edge deep learning models. Throughout your journey, you'll learn how to preprocess data, build and train models, evaluate performance, and deploy solutions, all within the MATLAB environment. Whether you're a beginner or an experienced practitioner, this course offers valuable insights and practical skills to excel in the rapidly evolving field of machine learning.

WHAT WE LEARNT

1. Introduction to MATLAB's machine learning capabilities and toolboxes.
2. Data preprocessing techniques such as normalization, feature extraction, and dimensionality reduction.
3. Supervised learning algorithms including regression (linear regression, support vector regression) and classification (k-nearest neighbors, decision trees, ensemble methods).
4. Unsupervised learning algorithms like clustering (k-means clustering, hierarchical clustering) and dimensionality reduction (principal component analysis, t-distributed stochastic neighbor embedding).
5. Evaluation metrics and techniques for assessing model performance.
6. Deep learning fundamentals, including neural network architectures and training strategies.
7. Practical skills for implementing machine learning and deep learning algorithms using MATLAB's built-in functions and toolboxes, as well as handling large datasets efficiently.

DURATION: 12hrs

GRADE: 100%

6. DEEP LEARNING ONRAMP

INTRODUCTION:

"Deep Learning Onramp" likely refers to a structured learning path or course designed to introduce individuals to the principles and applications of deep learning. Deep learning is a subset of machine learning that utilizes neural networks with multiple layers to learn intricate patterns and representations from data. This onramp may cover foundational concepts of deep learning, including neural network architecture, optimization techniques, convolutional neural networks (CNNs), recurrent neural networks (RNNs), and popular deep learning frameworks such as TensorFlow or PyTorch. Additionally, it might include hands-on exercises and projects to help learners gain practical experience in implementing deep learning models for various tasks such as image classification, natural language processing, and more.

WHAT WE LEARNT

1. Introduction to deep learning concepts, including neural networks, layers, and activation functions.
2. Building and training simple neural networks using frameworks like TensorFlow or PyTorch.
3. Convolutional Neural Networks (CNNs) for image recognition and classification tasks.
4. Recurrent Neural Networks (RNNs) for sequence modeling and prediction.
5. Transfer learning techniques for leveraging pre-trained models on new tasks.
6. Optimization algorithms and techniques for training deep learning models efficiently.
7. Model evaluation and deployment considerations.
8. Hands-on experience with practical projects and exercises to reinforce learning.

DURATION: 2hrs GRADE: 100%

7. DEEP LEARNING WITH MATLAB

INTRODUCTION:

"Deep Learning with MATLAB" refers to using MATLAB, a high-level programming language and environment, to implement and experiment with deep learning algorithms and models.

MATLAB provides a comprehensive set of tools and functions for deep learning, including prebuilt neural network architectures, optimization algorithms, and visualization tools. With MATLAB, users can easily preprocess data, design and train deep neural networks, fine-tune models, and deploy them for various applications such as image classification, object detection, speech recognition, and more. The integration of deep learning capabilities within MATLAB makes it a convenient and powerful platform for both beginners and experts in the field of deep learning.

WHAT WE LEARNT

1. Introduction to deep learning concepts, including neural networks, layers, and activation functions.
2. Building and training neural networks using MATLAB's Deep Learning Toolbox.
3. Convolutional Neural Networks (CNNs) for image classification, object detection, and segmentation tasks.
4. Recurrent Neural Networks (RNNs) for sequence modeling and prediction, such as in natural language processing or time series analysis.
5. Transfer learning techniques for fine-tuning pre-trained models on new datasets.
6. Optimization algorithms and techniques for training deep learning models efficiently.
7. Hyperparameter tuning and model evaluation strategies.
8. Deployment considerations, including exporting trained models for inference in production environments.
9. Practical projects and exercises to apply deep learning concepts to real-world datasets and problems.

DURATION: 7hrs

GRADE: 100%

8. COMPUTER VISION ONRAMP

INTRODUCTION

"Computer Vision Onramp" likely refers to a structured learning path or course designed to introduce individuals to the fundamentals of computer vision. Computer vision is a field of artificial intelligence that focuses on enabling machines to interpret and understand visual information from the real world. This onramp may cover foundational concepts such as image processing, feature extraction, object detection, image classification, and semantic segmentation. Additionally, it might include hands-on exercises and projects to help learners gain practical experience in implementing computer vision algorithms using popular frameworks like OpenCV or deep learning libraries such as TensorFlow or PyTorch.

WHAT WE LEARNT

1. Fundamentals of computer vision, including image representation and processing techniques.
2. Image filtering and feature extraction methods, such as edge detection and corner detection.
3. Object detection and recognition algorithms, including template matching and Haar cascades.
4. Image segmentation techniques for identifying and extracting objects from images.
5. Feature matching and registration methods for aligning images or detecting objects across multiple images.
6. Introduction to deep learning-based approaches for computer vision tasks, such as convolutional neural networks (CNNs) for image classification and object detection.
7. Practical skills for implementing computer vision algorithms using libraries like OpenCV or MATLAB's Computer Vision Toolbox
8. Hands-on projects and exercises to reinforce learning and apply computer vision techniques to real-world problems.

DURATION: 2hrs

GRADE:100%

9. REINFORCEMENT LEARNING ONRAMP

INTRODUCTION

"Reinforcement Learning Onramp" likely refers to a structured learning path or course designed to introduce individuals to the principles and applications of reinforcement learning (RL). Reinforcement learning is a type of machine learning paradigm where an agent learns to make decisions by interacting with an environment to maximize cumulative rewards. This onramp may cover foundational concepts such as Markov decision processes, policy evaluation, policy improvement, value iteration, Q-learning, and deep reinforcement learning. Additionally, it might include hands-on exercises and projects to help learners gain practical experience in implementing reinforcement learning algorithms for various tasks such as game playing, robotics, and autonomous systems.

WHAT WE LEARNT

1. Introduction to reinforcement learning (RL) concepts, including agents, environments, actions, and rewards.
2. Markov decision processes (MDPs) as a formal framework for modeling RL problems.
3. Dynamic programming methods such as value iteration and policy iteration for solving MDPs.
4. Model-free RL algorithms, including Q-learning and SARSA, for learning optimal policies directly from experience.
5. Deep reinforcement learning techniques, which combine deep learning with RL for handling high-dimensional state and action spaces.
6. Exploration-exploitation trade-offs and techniques for balancing exploration and exploitation in RL agents.
7. Function approximation methods for scaling RL algorithms to large state and action spaces.
8. Practical implementation of RL algorithms using libraries like TensorFlow, PyTorch, or RLlib.
9. Applications of reinforcement learning in various domains, such as robotics, game playing, and autonomous systems.

DURATION: 2.5hrs

GRADE: 100%

10.STATISTICS ONRAMP

INTRODUCTION

"Statistics Onramp" likely refers to a structured learning path or course designed to introduce individuals to the fundamentals of statistics. This onramp may cover essential statistical concepts such as descriptive statistics, probability theory, hypothesis testing, confidence intervals, regression analysis, and ANOVA (analysis of variance). Additionally, it might include hands-on exercises and projects to help learners understand how to apply statistical methods to analyze data, make informed decisions, and draw meaningful conclusions. Whether you're new to statistics or looking to refresh your knowledge, a Statistics Onramp can provide a solid foundation for understanding and working with data.

WHAT WE LEARNT

1. Fundamental concepts of descriptive statistics, including measures of central tendency (mean, median, mode) and variability (range, variance, standard deviation).
2. Probability theory, including basic probability rules and distributions such as the normal distribution, binomial distribution, and Poisson distribution.
3. Inferential statistics techniques for making inferences about populations based on sample data, including hypothesis testing and confidence intervals.
4. Understanding and interpreting p-values and confidence intervals in statistical analysis.
5. Regression analysis, including simple linear regression and multiple regression, for modeling relationships between variables.
6. Analysis of variance (ANOVA) for comparing means across multiple groups.
7. Basic concepts of experimental design and statistical power.
8. Practical skills for performing statistical analysis using software such as R, Python with libraries like NumPy and SciPy, or statistical software packages like SPSS or SAS.
9. Hands-on projects and exercises to apply statistical concepts to real-world datasets and problems.

DURATION: 1.5hrs

GRADE: 100%

11.MATLAB FOR DATA PROCESSING AND VISUALISATION

INTRODUCTION

"MATLAB for Data Processing and Visualization" refers to using MATLAB, a powerful programming language and environment, for tasks related to handling and visualizing data. In this context, MATLAB provides a wide range of tools and functions for processing, analyzing, and visualizing data in various formats, including numerical data, time series, images, and more. Users can leverage MATLAB's built-in functions for data manipulation, statistical analysis, filtering, interpolation, and other preprocessing tasks. Additionally, MATLAB offers sophisticated plotting and visualization capabilities, allowing users to create high-quality plots, charts, and interactive visualizations to explore and present their data effectively. Overall, MATLAB is a versatile tool for data scientists, engineers, and researchers to perform data processing and visualization tasks efficiently.

WHAT WE LEARNT

1. Data importing and exporting techniques to read data from various file formats such as CSV, Excel, and text files.
2. Data cleaning and preprocessing methods including handling missing data, removing outliers, and filtering noisy data.
3. Data manipulation and transformation techniques such as sorting, reshaping, and merging datasets.
4. Basic statistical analysis including calculating descriptive statistics, performing hypothesis tests, and fitting probability distributions.
5. Visualization techniques for exploring and presenting data, including plotting histograms, scatter plots, line plots, and bar graphs.
6. Customizing plots and adding annotations to enhance data visualization.
7. Interactive data exploration using MATLAB's graphical user interface (GUI) tools.
8. Advanced data visualization techniques such as 3D plotting, heatmap visualization, and geographic mapping.
9. Exporting visualizations for publication or presentation purposes

DURATION: 8 hrs

GRADE: 100%

12. IMAGE PROCESSING ONRAMP

INTRODUCTION

IMAGE PROCESSING ONRAMP is likely an introductory course or learning program focused on image processing. It could be offered by a university, online platform, or company to help individuals gain foundational knowledge and skills in image processing techniques, such as filtering, segmentation, feature extraction, and object recognition.

WHAT WE LEARNT

1. Fundamentals of digital image representation, including grayscale and color images.
2. Basic image processing operations such as image resizing, cropping, and rotation.
3. Image enhancement techniques including contrast adjustment, histogram equalization, and noise reduction.
4. Spatial domain filtering methods such as smoothing and sharpening filters.
5. Frequency domain techniques like Fourier Transform for image analysis and filtering.
6. Edge detection algorithms such as Sobel, Prewitt, and Canny for detecting edges in images.
7. Image segmentation methods for partitioning an image into meaningful regions.
8. Feature extraction techniques for extracting relevant information from images.
9. Introduction to computer vision concepts and applications.
10. Practical skills for implementing image processing algorithms using MATLAB's Image Processing Toolbox or libraries like OpenCV.
11. Hands-on projects and exercises to apply image processing techniques to real-world image datasets and problems.

DURATION: 2hrs

GRADE: 100%

13.IMAGE PROCESSING WITH MATLAB

INTRODUCTION

"Image Processing with MATLAB" is a course or a set of resources designed to teach image processing techniques using the MATLAB software. MATLAB is a programming environment and language commonly used in engineering and scientific fields, including image processing. This course likely covers topics such as image enhancement, filtering, segmentation, feature extraction, and object recognition, all using MATLAB's built-in functions and toolboxes. It's a practical way to learn how to apply image processing algorithms using MATLAB's powerful tools.

WHAT WE LEARNT

1. Basic image representation and manipulation techniques, including loading, displaying, and saving images.
2. Image enhancement methods such as contrast adjustment, histogram equalization, and noise reduction.
3. Spatial domain filtering techniques for smoothing and sharpening images using various filter kernels.
4. Frequency domain analysis using Fourier Transform for image filtering and analysis.
5. Edge detection algorithms like Sobel, Prewitt, and Canny for detecting edges in images.
6. Image segmentation techniques for partitioning images into meaningful regions.
7. Morphological operations such as dilation, erosion, opening, and closing for shape analysis and feature extraction.
8. Feature extraction methods for extracting relevant information from images, such as texture analysis and object recognition.
9. Introduction to advanced topics such as image registration, image stitching, and image restoration.

10. Practical implementation of image processing algorithms using MATLAB's Image Processing Toolbox.

11. Hands-on projects and exercises to apply image processing techniques to real-world image datasets and problems.

DURATION : 11hrs

GRADE: 100%

14.SIGNAL PROCESSING WITH MATLAB

INTRODUCTION

"Signal Processing with MATLAB" is a course or educational resource focused on teaching signal processing techniques using the MATLAB software. Signal processing involves analyzing, manipulating, and interpreting signals, which could be audio, video, or any other type of data that varies over time or space. MATLAB provides a rich set of functions and toolboxes specifically designed for signal processing tasks, making it a popular choice for engineers, researchers, and students working in this field. This course likely covers topics such as filtering, spectral analysis, time-frequency analysis, and digital signal processing algorithms, all implemented using MATLAB.

WHAT WE LEARNT

1. Basics of signal representation and analysis, including continuous-time and discrete-time signals.
2. Time-domain and frequency-domain analysis techniques, such as Fourier Transform, Discrete Fourier Transform (DFT), and Fast Fourier Transform (FFT).
3. Filtering operations, including design and implementation of digital filters such as low-pass, high-pass, band-pass, and band-stop filters.
4. Convolution and correlation operations for signal processing tasks.
5. Windowing techniques for signal analysis and spectral estimation.
6. Spectral analysis methods, including periodogram, power spectral density (PSD), and spectrogram.
7. Introduction to adaptive filtering techniques for applications such as noise cancellation and system identification.

8. Practical implementation of signal processing algorithms using MATLAB's Signal Processing Toolbox.

9. Hands-on projects and exercises to apply signal processing techniques to real-world signal datasets and problems.

DURATION: 7.5hrs

GRADE: 100%

15.WIRELESS COMMUNICATION ONRAMP

INTRODUCTION

"Wireless Communication ONRAMP" sounds like a specialized learning program or course designed to introduce individuals to the fundamentals of wireless communication technologies. This could include topics such as radio frequency (RF) communication, modulation techniques, multiple access schemes, channel coding, and wireless networking protocols. The term "ONRAMP" suggests that it's an entry-level or introductory program, aimed at providing a foundational understanding of wireless communication concepts. It could be offered by educational institutions, online platforms, or companies as a means for individuals to gain essential knowledge in this field.

WHAT WE LEARNT

1. Introduction to wireless communication onramp
2. How they are simulate a basic digital communication link ,anatomy of digital communication system and modulation and demodulation
3. Simulate a digital filters
4. Multipath channels
5. ODFM modulation and practical consideration anad null subcarriers

DURATION: 1 HOUR

GRADE: 100%

16.OPTIMIZATION ONRAMP

INTRODUCTION

"Optimization ONRAMP" likely refers to an educational program or course focused on optimization techniques. Optimization is a branch of mathematics and computer science that deals with finding the best solution to a problem from a set of possible solutions. This could involve maximizing or minimizing an objective function while satisfying certain constraints. The term "ONRAMP" suggests that it's an introductory or beginner-level program aimed at helping individuals learn the basics of optimization methods and algorithms. It could cover topics such as linear programming, nonlinear optimization, convex optimization, and heuristic optimization techniques.

DURATION:3hrs

GRADE: 100%

17.MATLAB PROGRAMMING TECHNIQUES

INTRODUCTION

"MATLAB Programming Techniques" refers to strategies, methods, and best practices for writing efficient and effective MATLAB code. This could include topics such as:

1. Syntax and language features: Understanding MATLAB syntax, data types, and built-in functions.
2. Vectorization: Leveraging MATLAB's ability to perform operations on entire arrays instead of individual elements for faster execution.
3. Memory management: Writing code that minimizes memory usage and optimizes memory access patterns.
4. Code organization: Structuring MATLAB code into functions, scripts, and classes for better readability and maintainability.
5. Debugging and error handling: Techniques for identifying and fixing errors in MATLAB code, as well as handling exceptions gracefully.
6. Performance optimization: Strategies for improving the performance of MATLAB code, such as preallocation, avoiding unnecessary computations, and using efficient algorithms.
7. MATLAB toolboxes: Utilizing built-in MATLAB toolboxes and libraries to streamline development and leverage specialized functionality.

These techniques help MATLAB programmers write code that is not only correct but also efficient, scalable, and easy to maintain.

WHAT WE LEARNT

1. Basic syntax and programming constructs in MATLAB, including variables, data types, operators, and control flow statements (if-else, loops).
2. Writing modular and reusable code using functions and scripts.
3. Vectorization techniques for optimizing code performance and efficiency.
4. Debugging and error handling strategies to identify and fix issues in MATLAB code.
5. MATLAB programming best practices, including code readability, documentation, and naming conventions.
6. Advanced data structures such as cell arrays, structures, and tables for organizing and manipulating data.
7. File I/O operations for reading and writing data to files in various formats (e.g., text files, Excel files).
8. Object-oriented programming (OOP) concepts in MATLAB, including classes, objects, inheritance, and polymorphism.
9. MATLAB's built-in libraries and toolboxes for specialized tasks such as signal processing, image processing, and optimization.
10. Profiling and optimization techniques for improving code performance and efficiency.
11. Version control and collaborative coding practices using MATLAB's integration with Git and GitHub.

DURATION: 16hrs

GRADE: 100%

18.SOLVING NON-LINEAR EQUATIONS WITH MATLAB

INTRODUCTION:

Solving nonlinear equations with MATLAB involves using numerical methods to find approximate solutions to equations that cannot be solved analytically. MATLAB provides several built-in functions and tools for solving nonlinear equations, including:

1. `fsolve`: This function finds a solution to a system of nonlinear equations. It requires an initial guess for the solution and returns the solution that satisfies the equations.
2. `fzero`: This function finds a root of a single-

variable nonlinear function. It requires an initial guess for the root and returns the value of the root. 3. Nonlinear equation solvers in optimization toolboxes: MATLAB's optimization toolboxes provide additional functions for solving nonlinear equations, including `lsqnonlin` for least squares problems and `lsqcurvefit` for curve fitting. To solve a nonlinear equation in MATLAB, you typically define the equation as a function or anonymous function, specify initial guesses for the solution, and then use one of the above functions to find the solution. MATLAB's documentation provides detailed information and examples for using these functions to solve nonlinear equations.

WHAT WE LEARNT

In a course on solving nonlinear equations with MATLAB, you typically learn:

1. Introduction to nonlinear equations and their importance in various fields such as engineering, physics, and economics.
2. Different methods for solving nonlinear equations, including numerical methods like the Newton-Raphson method, the secant method, and the bisection method.
3. Understanding the principles behind each method, including their convergence properties and limitations.
4. Implementing these methods in MATLAB using built-in functions or custom scripts.
5. Handling complex nonlinear systems of equations using iterative methods and numerical techniques.
6. Evaluating the accuracy and efficiency of different solution methods for nonlinear equations.
7. Strategies for dealing with singularities, multiple solutions, and other challenges that may arise in nonlinear equation solving.
8. Applications of nonlinear equation solving in practical problems, such as root-finding in optimization and parameter estimation tasks. Overall, the course aims to equip students with the knowledge and skills necessary to effectively solve nonlinear equations using MATLAB, enabling them to tackle a wide range of mathematical and engineering problems.

DURATION: 3hrs

GRADE: 100%

MATALAB CERTIFICATE

MATALAB Onramp



Course Completion Certificate

Shyam A

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

MATLAB Onramp

A handwritten signature in black ink, reading 'Craig Santos'.
DIRECTOR, TRAINING SERVICES

13 November 2023

Object Oriented Programming Onramp



Course Completion Certificate

Shyam A

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

Object-Oriented Programming Onramp

A handwritten signature in black ink, reading 'Craig Santos'.
DIRECTOR, TRAINING SERVICES

14 December 2023



Course Completion Certificate

Shyam A

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

Machine Learning Onramp



DIRECTOR, TRAINING SERVICES

15 December 2023



Course Completion Certificate

Shyam A

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

Computer Vision Onramp



DIRECTOR, TRAINING SERVICES

16 December 2023



Course Completion Certificate

Shyam A

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

Reinforcement Learning Onramp



DIRECTOR, TRAINING SERVICES

17 December 2023

MATLAB for Data Processing and Visualization



Course Completion Certificate

Shyam A

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

MATLAB for Data Processing and Visualization



DIRECTOR, TRAINING SERVICES

2 February 2024

Deep Learning with MATLAB



Course Completion Certificate

Shyam A

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

Deep Learning with MATLAB

A handwritten signature in black ink, reading "Craig Santos".

DIRECTOR, TRAINING SERVICES

9 February 2024

Machine Learning with MATLAB



Course Completion Certificate

Shyam A

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

Machine Learning with MATLAB

A handwritten signature in black ink, reading "Craig Santos".

DIRECTOR, TRAINING SERVICES

12 February 2024



Course Completion Certificate

Shyam A

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

Deep Learning Onramp

A handwritten signature in black ink, reading "Craig Santos".

DIRECTOR, TRAINING SERVICES

12 February 2024

MATLAB Programming Techniques



Course Completion Certificate

Shyam A

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

MATLAB Programming Techniques

A handwritten signature in black ink, reading "Craig Santos".

DIRECTOR, TRAINING SERVICES

17 February 2024



Course Completion Certificate

Shyam A

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

Statistics Onramp



DIRECTOR, TRAINING SERVICES

1 March 2024

Image Processing Onramp



Course Completion Certificate

Shyam A

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

Image Processing Onramp



DIRECTOR, TRAINING SERVICES

5 March 2024



Course Completion Certificate

Shyam A

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

Solving Nonlinear Equations with MATLAB



DIRECTOR, TRAINING SERVICES

5 March 2024

Signal Processing Onramp



Course Completion Certificate

Shyam A

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

Signal Processing Onramp



DIRECTOR, TRAINING SERVICES

7 March 2024



Course Completion Certificate

Shyam A

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

Simulink Onramp



DIRECTOR, TRAINING SERVICES

28 March 2024

Image Processing with MATLAB



Course Completion Certificate

Shyam A

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

Image Processing with MATLAB



DIRECTOR, TRAINING SERVICES

5 April 2024



Course Completion Certificate

Shyam A

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

Signal Processing with MATLAB

DIRECTOR, TRAINING SERVICES

6 April 2024

MATLAB FUNDAMENTALS



Course Completion Certificate

Shyam A

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

MATLAB Fundamentals

DIRECTOR, TRAINING SERVICES

4 February 2024