

Integrated analysis: Gene Expression, Survival, Biomarker, and Drug Target

Submitted in partial fulfilment of the required for the award of
Master of Science Degree in Bioinformatics and Data Science

By **M.P.KEERTHANA (43739041)**



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
GANDHISALAI, CHENNAI – 600119**

APRIL 2024



SATHYABAMA

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 M.P.KEERTHANA (43739041) who carried out the project entitled “Integrated Analysis: Gene Expression, Survival, Biomarkers, and Drug Target” under Dr. JEMMY CHRISTY 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, M.P.KEERTHANA (**43739041**), hereby declare that the Project Report “Integrated Analysis: Gene Expression, Survival, Biomarkers, and Drug Target” done by me under the guidance of Dr. JEMMY CHRISTY is submitted in partial fulfilment of the requirements for the award of Master of Science degree in Bioinformatics and Data Science

DATE:

M.P.KEERTHANA

PLACE: CHENNAI

SIGNATURE OF THE CANDIDATE

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to 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 sence of gratitude to my project guide Dr. JEMMY CHRISTY Department of Bioinformatics and Data Science for his valuable guidance, suggestions and constant encouragement 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.

(M.P.KEERTHANA)

INDEX

S.NO	TITLE	PAGE NO
1	TRANSLATIONAL BIOINFORMATICS	6
1.1	GENE EXPRESSION ANALYSIS	6
1.2	CUMULATIVE SURVIVAL PROBABILITY CALCULATION USING KM PLOTTER	9
1.3	BIOMARKERS: IDENTIFYING BIOLOGICAL MARKERS FOR DISEASE DIAGNOSIS AND THERAPEUTIC MONITORING	13
1.4	DRUG TARGET INTERACTION PREDICTION: APPROACHES FOR PREDICTING INTERACTIONS SPECIFIC TO PATIENT TREATMENT AND DRUG DESIGN	15
2	PLAGIARISM REPORT	20
2.1	REVIEW ARTICLE	21
3	MATLAB COURSES	26
3.1	MATLAB CERTIFICATES	34

1. TRANSLATIONAL BIOINFORMATICS

1.1 GENE EXPRESSION ANALYSIS

INTRODUCTION:

A biological network is a method of representing systems as complex sets of binary interactions or relations between various biological entities. A typical graphing representation consists of a set of nodes connected by edges. Cytoscape is an open source software platform for visualising molecule interactions network and biological pathway and integrating the network with annotations, gene expression profile. Cytohubba predicts and explores the important nodes and edges in a given network by several topological algorithms such as Degree, Bottleneck, MCC, DMNC, EPC, Eccentricity, Closeness, Stress, Clustering coefficient and Betweenness.

AIM:

Build a network to understand gene interrelationship and regulatory mechanism.

MATERIALS AND METHODS:

- GEO2R is an interactive web tool that allows users to compare two or more groups of Samples in a GEO Series in order to identify genes that are differentially expressed across experimental conditions.
- CYTOSCAPE: Cytoscape is an open source software platform for visualising molecule interactions network and biological pathway and integrating the network with annotations, gene expression profile.
- CYTOHUBBA: Cytohubba provides 11 topological analysis methods including Degree, Edge Percolated Component, Maximum Neighborhood Component, Density of Maximum Neighborhood Component, Maximal Clique Centrality and six centralities (Bottleneck, Eccentricity, Closeness, Radiality, Betweenness, and Stress) based on shortest paths.
- CLUEGO: Cluego is a Cytoscape App that facilitates the biological interpretation of large lists of genes and proteins by selecting representative Gene Ontology terms and pathways from multiple ontologies and visualises them into functionally organised networks.

PROCEDURE:

- For the database on human kidney cancer GEO2R is the source of the geo dataset (GSE66272).

- After choosing the suitable parameters, the GEO dataset is retrieved.
- Genes that were upregulated were copied and paste into the cytoscape network.
- Build a biological network using a BioGrid in cytohubba to analyse the network and identify importance node based various topologies algorithm to determine the top 10 genes.
- Genes must be copied in cluego and continue cluego parameter such as species ontology and statistical settings.
- Run cluego analysis to identify functional categories or pathways associated with the selected nodes.

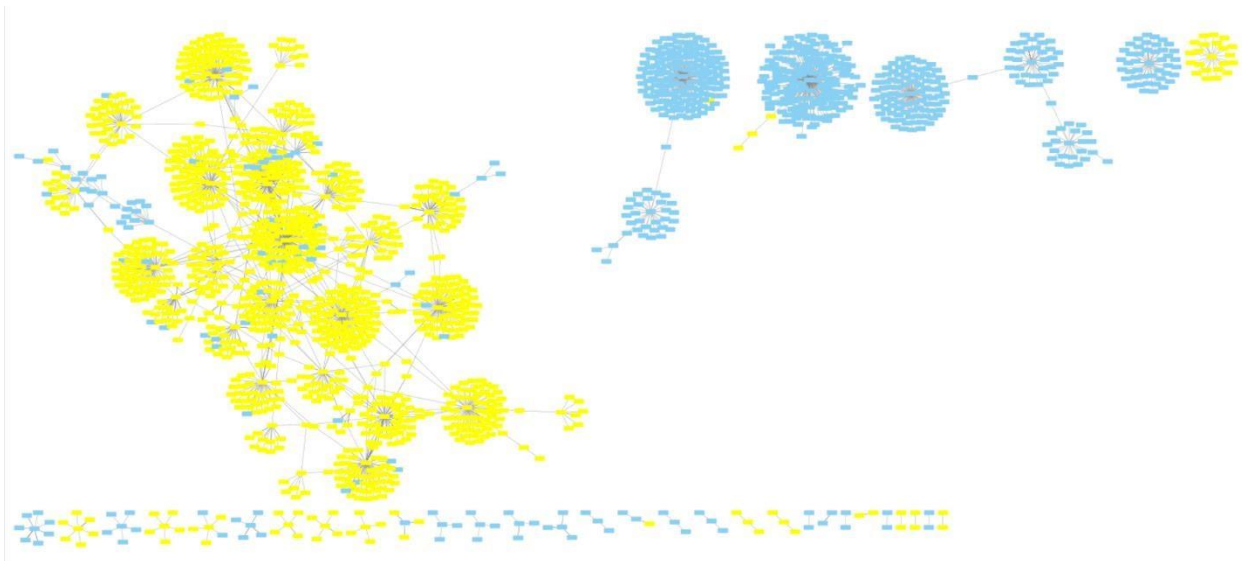


Figure 1: Network pathway of upregulated gene

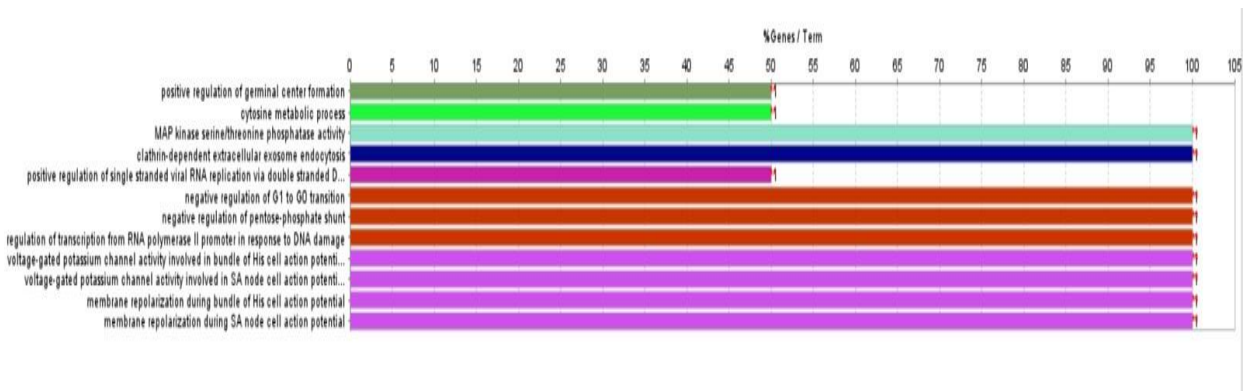


Figure 2: Graph of functional analysis

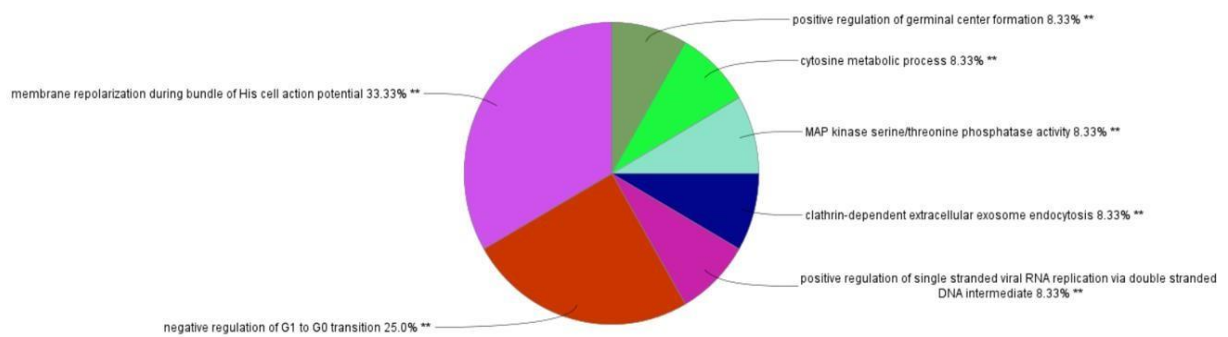


Figure 3: Pie chart of functional analysis

INTERPRETATION:

GEO Dataset GSE ID : GSE66272

Sample no: 54

Platform: GPL570

Type: Expression profiling by array

No of differentially expressed genes: 54676

No of upregulated genes: 64

No of downregulated genes: 54612

Network features of upregulated genes: nodes- 1808 , edges- 2347

1.2. 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.

To avoid missing correlations due to the use of a specific cutoff, all available cutoff values between the lower and upper quartiles of expression are used for the selected gene, and false discovery rate (FDR) using the Benjamini-Hochberg method is computed to correct for multiple hypothesis testing. The cutoff value with the highest significance (lowest FDR) is determined. In case of multiple cutoff values with identical significance, the cutoff with the highest hazard (HR) rate is selected for the final analysis.

AIM:

Calculate Proposed gene TOP2A cumulative survival probabilities.

INTERPRETATION

- KM-plot recognizes proposed gene symbols (including HUGO Gene Nomenclature Committee approved)
- Sources for the database include GEO, EGA, TCGA, Metabric, Impact, and PubMed repositories

- **BLADDER CARCINOMA:**

HR Rate: 1.23(0.89-1.68)

50% of median survival of probability:

Low expression cohort: 38.77

High expression cohort: 26.5

- **BREAST CANCER:**

HR Rate: 0.81(0.59-1.12)

50% of median survival of probability:

Low expression cohort: 122.3

High expression cohort: 148.53

-

CERVICAL SQUAMOUS CELL CARCINOMA

HR Rate: 0.69(0.42-1.11)

50% of median survival of probability:

Low expression cohort: 96.27

High expression cohort: 136.2

- ESOPHAGEAL ADENOCARCINOMA:

HR Rate: 1.8(0.92-3.5)

50% of median survival of probability:

Low expression cohort: 46.73

High expression cohort: 14.5

- ESOPHAGEAL SQUAMOUS CELL CARCINOMA:

HR Rate: 0.31(0.11-0.86)

50% of median survival of probability:

Low expression cohort: 25.43

High expression cohort: 48.6

- HEAD NECK SQUAMOUS CELL CARCINOMA:

HR Rate: 1.22(0.9-1.65)

50% of median survival of probability:

Low expression cohort: 68.8

High expression cohort: 48.63

- KIDNEY RENAL CELL CARCINOMA:

HR Rate: 5.37(2.94-9.82)

50% of median survival of probability:

Low expression cohort: 89.47

High expression cohort: 20.8

- LIVER HEPATOCELLULAR CARCINOMA:

HR Rate: 2.14(1.51-3.04)

50% of median survival of probability:

Low expression cohort: 71.03

High expression cohort: 25.6

- LUNG ADENOCARCINOMA:

HR Rate: 1.84(1.34-2.52)

50% of median survival of probability:

Low expression cohort: 59.67

High expression cohort: 40.97

- LUNG SQUAMOUS CELL CARCINOMA:

-

HR Rate: 0.67(0.51-0.88)

50% of median survival of probability:

Low expression cohort: 35.83

High expression cohort: 63.73

OVARIAN CANCER:

HR Rate: 0.9(0.68-1.18)

50% of median survival of probability:

Low expression cohort: 45.47

High expression cohort: 44.13

- PANCREATIC DUCTAL ADENOCARCINOMA:

HR Rate: 2.49(1.58-3.93)

50% of median survival of probability:

Low expression cohort: 37.67

High expression cohort: 16.6

- PHEOCHROMOCYTOMA & PARAGANGLIOMA:

HR Rate: 4.59(0.81-25.91)

50% of median survival of probability:

Low expression cohort: N/A

High expression cohort: N/A

- RECTUM ADENOCARCINOMA:

HR Rate: 0.43(0.19-0.96)

50% of median survival of probability:

Low expression cohort: 36.53

High expression cohort: 43.8

- SARCOMA:

HR Rate: 2.2(1.39-3.47)

50% of median survival of probability:

Low expression cohort: 47.47

High expression cohort: 20.93

- STOMACH ADENOCARCINOMA:

HR Rate: 0.71(0.45-1.03)

50% of median survival of probability:

Low expression cohort: 26.47

High expression cohort: 56.2

- TESTICULAR:

HR Rate: 0(0-Inf)

- 50% of median survival of probability:
Low expression cohort: N/A
High expression cohort: N/A

- THYMOMA:
HR Rate: 0.12(0.02-0.58)
50% of median survival of probability:
Low expression cohort: N/A
High expression cohort: N/A

THYROID CARCINOMA:
HR Rate: 1.92(0.67-5.54)
50% of median survival of probability:
Low expression cohort: N/A
High expression cohort: N/A

- UTERINE CORPUS ENDOMETRIAL CARCINOMA:
HR Rate: 1.76(1.14-2.72)
50% of median survival of probability:
Low expression cohort: 108.37
High expression cohort: 41.37

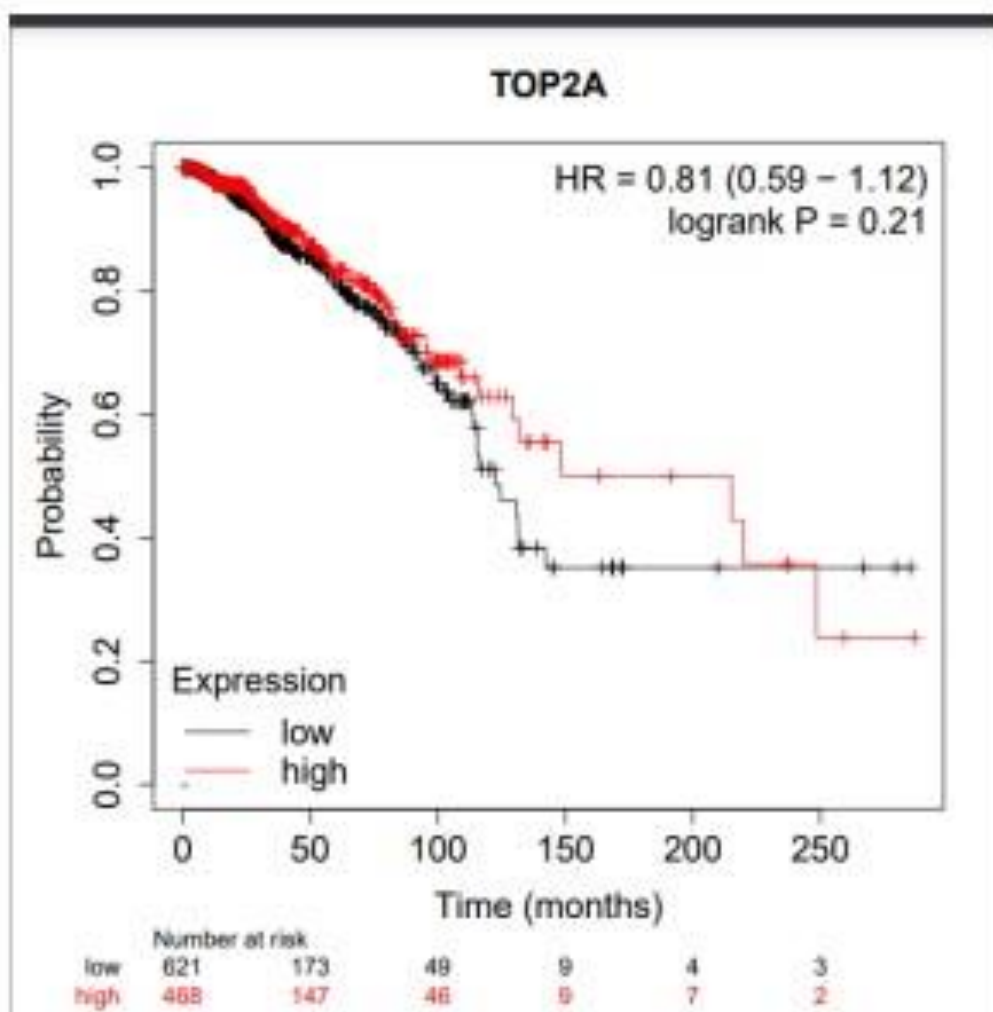


Figure 4 : BREAST CANCER

1.3. BIOMARKERS: IDENTIFYING BIOLOGICAL MARKERS FOR DISEASE DIAGNOSIS AND THERAPEUTIC MONITORING

DESCRIPTION:

TheMarker (a) systematically offers five types of therapeutic biomarkers, which realize 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 TheMarker were expected to have great implications and impact on the clinical development and clinical practice of drugs.

AIM:

Identify the disease markers related to Alzheimer disease 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 Alzheimer disease data's were retrieved and literature survey was done

INTERPRETATION:

GENERAL INFORMATION OF THE BIOMARKER

BIOMARKER ID	BIOMARKER NAME	CLASS	GENE NAME	BIOMARKER TYPE	REPRESENTATIVE DRUG	REPRESENTATIVE DISEASE
BQ2V5N	Amyloid beta A4 protein	protein	APP; A4; ADI	PDY MOI PRD	BACE1 inhibitor	Alzheimer disease

BIOMARKER TYPE AND RELAVANT DATA:

- Predictive biomarker (PRD): 3
- Safety biomarker (SAF): 5

- Pharmacodynamic biomarker (PDY): 25
- Monitoring biomarker (MOI): 2
- Surrogate biomarker (SUR): 1

SAFETY BIOMARKER(SAF):

Application of the Biomarker in Drug Development and Clinical Practice

Drug Status: Approved Drug

Drug Name: Donepezil

Drug Status: Approved Drug

Drug Name: Galatamine

Drug Status: Approved Drug

Drug Name: Rivastigmine

Drug Status: Approved Drug

Drug Name: Etanercept

Drug Status: Approved Drug Drug

Name: Infliximab

APPLICATION OF THE BIOMARKER IN DIAGNOSIS AND PROGNOSIS:

DIAGNOSIS:

- Alzheimer disease
- Breast cancer

TISSUE-SPECIFIC EXPRESSION OF THE BIOMARKER:

- Tissue-Specific RNA Expression of this Biomarker in total 32 tissues
- Tissue-Specific Protein Abundance of this Biomarker in total 32 tissues

1.4. 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.
- **SwissTarget Prediction:** SwissTargetPrediction 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.
- SwissTargetPrediction 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 SwissTargetPrediction, 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 Donepezil targets screening and binding assessment.

MATERIAL AND METHODS:

SWISS TARGET PREDICTION

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.

PDB: The Protein Data Bank is a repository for the 3-D structural data of large biological molecules, such as proteins and nucleic acid. The PDB format accordingly provides for description and annotation of protein and nucleic acid structures including atomic coordinates, secondary structure assignments, as well as atomic connectivity.

DRUGBANK: DrugBank Online is a comprehensive, free-to-access, online database containing information on drugs and drug targets. As both a bioinformatics and a cheminformatics resource, we

combine detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information.

PROCEDURE:

- Disease/Disorder Alzheimer associated FDA approved drugs was chosen and smiled notation was retrieved from Pubchem using following link
<https://pubchem.ncbi.nlm.nih.gov/>
- Swisstarget home page was accessed using following link
<http://www.swisstargetprediction.ch/index.php>
- Predicted the targets (Top five) and drug interaction was analyzed using discovery studio Receptor ligand interaction module

INTERPRETATION:

DRUG NAME: Donepezil

NO.OF. TARGET PREDICTED: 65

CHOOSEN TARGET: 2

HUB GENES SELECTED: HSP90N, p33, YAL021C, HEL-S-77p, RIPK7, RNF53, SPCC5E4.02c, UBC, v-abl, FRAP1

UNIPROT ID	PDB ID	NAME OF THE PROTEIN	RESOLUTION VALUE	CHAIN	DOMAIN
Q12931	1HOO	Heat shock protein 75 KDa	2.30A ⁰	A and B	1-52

P42345	4DRI	Serine / threonineprotein kinase mTOR	1.45A ⁰	B	1-72
--------	------	--	--------------------	---	------

TARGET	COMMON NAME	UNIPROT ID	TARGET CLASS	PROBABILITY
Butyrl cholinestrse	BCHE	P06276	Hydrolase	0.97150344
Acetyl cholinesterase	ACHE	P22303	Hydrolase	0.97150344
HERG	KINM2	Q12309	Voltage-gated ion channel	0.97150344
Histamine H3 receptor	HRH3	Q9Y5NI	Family AG protein coupled receptor	0.97150344
Sigma opioid receptor	SIGMAR1	Q99720	Membrane receptor	0.97150344

MOLECULE TABLE:

NAME	-PLP1	-PLP2	JAIN	PMF	DOCK SCORE
1HOO	46.09	46.01	1.02	87.14	7.759
3152					
4DRI	39.87	44.56	-0.74	112.35	60.431
3152					

NON-BOND TABLE:

NAME	DISTANCE	CATEGORY	TYPE	FROM	TO
1HOO	2.62482	Hydrogen bond	Conventional hydrogen bond	A:ASN132	3152:01
A:ASN132:ND2					
4DRI	2.38102				
A:LYS121:ND		Hydrogen bonding	Conventional hydrogen bond	A:LYS121	3152:02

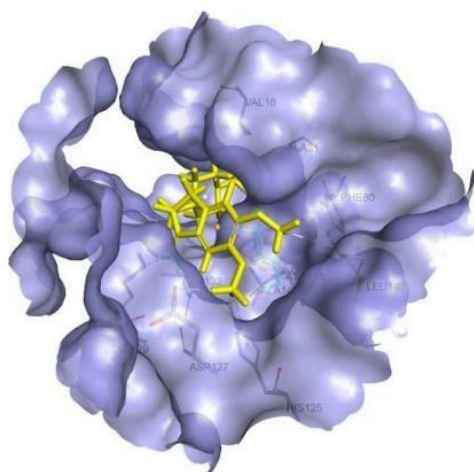


Figure 5.1: 3D STRUCTURE OF 1HOO

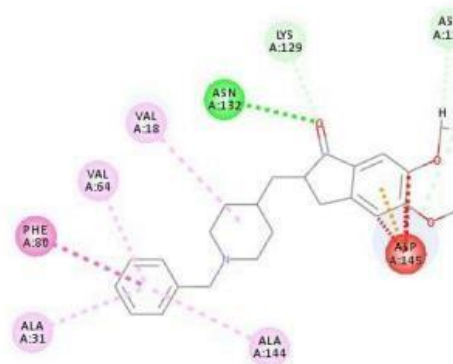


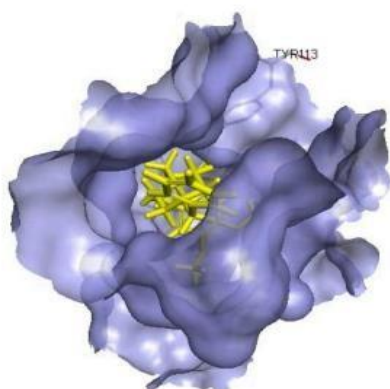
Figure 5.2: 2D STRUCTURE OF 1HOO

ASN132 RESIDUE INTERACT WITH 1HOO:

ASN132 residue with other ligands or hydrogen bonding within the protein is significant for the protein structure and function. It involved in stabilizing the protein structure, mediating proteinprotein interaction are participating in ligand bonding. ASN132 residue and the 1HOO protein can be used in structure based on drug design.

VAL18 RESIDUE INTERACT WITH 1HOO:

The interaction of the VAL18 residue with the 1HOO protein could also be significant in drug discovery although its importance would depend on its location within the protein structure and its role in protein function. VAL18 may contribute to the structural stability of the protein.



4DRI

Figure 6.1: 3D STRUCTURE OF 4DRI

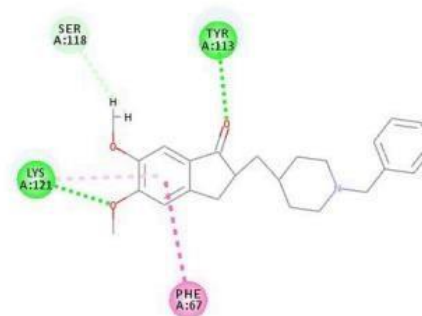


Figure 6.2: 2D STRUCTURE OF

SER118 RESIDUE INTERACT WITH 4DRI:

SER118 might be located in a structurally important region of the protein. It could be involved in stabilizing SER118 and its neighbouring residues can help in target identification and validation for drug discovery.

LYS121 RESIDUE INTERACT WITH 4DRI:

LYS121 residue interacts with 4DRI protein in drug discovery. LYS121 might be located in a structurally important region of the network, possibly involved in stabilizing the protein structure through interaction with other residues or ligands.

3. REVIEW ARTICLE

M.P.Keerthana Murali

ORIGINALITY REPORT

3%

SIMILARITY INDEX

1%

INTERNET SOURCES

2%

PUBLICATIONS

0%

STUDENT PAPERS

PRIMARY SOURCES

1

Heike E Assmus, Ralf Herwig, Kwang-Hyun Cho, Olaf Wolkenhauer. "Dynamics of biological systems: role of systems biology in medical research", Expert Review of Molecular Diagnostics, 2014

Publication

2%

2

www.frontiersin.org

Internet Source

1%

Exclude quotes Off

Exclude matches Off

Exclude bibliography On

FROM GENES TO SYSTEMS: A COMPREHENSIVE REVIEW OF SYSTEMS BIOLOGY

GENOME SCALE MODELLING AND SIMULATION IN SYSTEMS BIOLOGY:

Genome-scale modelling and simulation stand as pivotal methodologies in the domain of systems biology, offering a holistic perspective on the intricacies of biological systems. In this comprehensive review, we delve into the fundamental principles, methodologies, and applications of genome-scale modelling and simulation. At its core, genome-scale modelling encompasses the reconstruction and analysis of metabolic, regulatory, and signalling networks, spanning from the subcellular to the cellular level. These models integrate diverse omics data, including genomics, transcriptomics, proteomics, and metabolomics, to elucidate the complex interactions among biological components.

Genome-scale modeling utilizes constraint-based modeling, which features flux balance analysis and constraint-based reconstruction and analysis (COBRA), for analyzing metabolic networks in a variety of computational approaches. Dynamic modeling approaches that reflect the spatial dynamics of regulatory and signaling networks include stochastic simulations and ordinary differential equations (ODEs). Also, hybrid methods of modeling integrate the advantages of dynamic and constraint-based modeling for comprehensive system analysis.

Genome-scale modeling and simulation are made simpler by a variety of software tools, which include SimBiology, BioNetGen, PySCeS, Systems Biology Markup Language (SBML), CellDesigner, COBRA Toolbox, and COPASI. These tools provide places for the growth, development, and analysis of models, allowing researchers to investigate complex biological processes.

Genome-scale modeling provides a wide range of significant applications. By improving biological pathways, metabolic engineering utilizes these models for creating technological advancement and biofuel production by optimizing cellular pathways. Genome-scale models enhance target identification, therapeutic repurposing, and toxicity prediction in the development and discovery of drugs. In addition, patient-specific modeling facilitates precision medicine and therapeutic optimization in personalized healthcare.

However, there are some challenges with genome-scale modeling. Prediction reliability and accuracy are still greatly reliant on parameter estimation and validation of models. Computational models and experimental data integration are complicated processes that require reliable techniques for both data integration and improving the model. Additionally, maintaining large-scale biological networks poses challenges because of scalability and complexity constraints, and requires advancements in hardware infrastructure and computational techniques. Systems biology promises to offer a bright future for genome-scale modeling and simulation. Developments in multiscale modeling and its integration with other modeling techniques will improve our understanding of the dynamics and behavior of biological systems. In the end, genome-scale modeling is a potent tool for deciphering life's intricacies and propelling advances in biotechnology, medicine, and other fields.

SYSTEM BIOLOGY APPROACHES FOR STUDYING BIOLOGICAL DYNAMICS:

Systems in biology are described as collections of connected and interacting things that appear at several levels of organization, such as the individual, population, tissue, organ, subcellular, and cellular levels. Understanding is the primary objective of systems analysis. The characteristics of dynamic processes result in the combination of dynamic systems theory and biology to form systems biology. The goal of this multidisciplinary field is to gain insight into the dynamic interactions and resulting functional activities of biological components rather than simply identifying and characterizing them.

The dynamic aspect of biological processes, including gene expression, metabolism, and signaling, is highlighted by systems biology. For example, cell differentiation is a dynamic, nonlinear process that is impacted by varying intensity levels and durations of stimuli. Rather cell differentiation is identified as a bistable system using systems-theoretical analysis, which enables optimized experimental designs through simulations rather than depending just on knowledge of molecular constituents and static representations.

Because of interactions between system components, biological systems show dynamic changes that explain development, adaptation to environmental changes, and the course of disease. As a result, the subject matter needs to change from static to perspectives on biological systems that are dynamic. Proteome, transcriptome, and metabolome analysis on a massive scale have been made possible by recent developments in high-throughput technology. These 'omics' data, however, provide static descriptions of systems. It is essential to take several snapshots at various times to reconstruct system dynamics and to systematically couple high-throughput data collection with experimental perturbations to observe the changes that arise.

Systems biology, in its simplest form, is the integration of dynamic systems theory with biology. It highlights the dynamic character of biological processes and the significance of understanding the interactions between system components to fully explain system behavior and function.

SYSTEM BIOLOGY INSIGHTS INTO EVOLUTIONARY PROCESSES AND ADAPTATION:

Evolutionary biology has long contributed to our understanding of life diversity; systems biology has improved this understanding by offering new insights into the fundamental mechanisms driving evolutionary processes and adaptability. This looks at how methods from systems biology help us understand evolution and adaptation at various biological scales. Within a systems biology paradigm, we examine how functional genomics, network analysis, quantitative modeling, comparative studies, and genomics combine to reveal the intricate processes of evolution. We illustrate the complementary nature of systems biology and evolutionary biology by drawing on knowledge from a broad range of disciplines. This helps to clarify the molecular, cellular, and ecological factors that affect an organism's evolutionary trajectory.

Comparative genomics provides information on evolutionary conservation in the field of genomics and evolution by identifying genetic features that are shared by species and by revealing genomic signals that are indicative of adaptation and selection. By revealing the evolutionary limitations found in protein-protein interaction networks and the dynamics of metabolic and gene regulatory networks throughout evolutionary time scales, network analysis within an evolutionary framework explores the complex structure of biological systems. Using techniques like population genetics models and evolutionary game theory, quantitative modeling offers a predictive framework for comprehending evolutionary dynamics. Paths of adaptation within a population. The field of functional genomics investigates the molecular mechanisms behind adaptive responses, including transcriptomics, proteomics, and metabolomics. This research provides a spotlight on the molecular basis of evolutionary change. Ecological interactions are essential to evolution; co-evolutionary processes and the topologies of ecological networks are shaped by symbiotic partnerships and hostpathogen interactions. Using comparative genomes and phylogenetics, comparative systems biology expands our understanding of other species and reveals patterns of adaptive convergence and convergent evolution.

INTERDISCIPLINARY PERSPECTIVES ON SYSTEM BIOLOGY RESEARCH:

This review article provides an in-depth investigation of how systems biology applies knowledge and methods from various scientific disciplines to unravel the complex mechanisms of biological systems. By looking at the contributions of the fields of philosophy, ecology, biology, physics, chemistry, engineering, bioinformatics, and mathematics, we provide insight into the multidisciplinary nature of systems biology. Additionally, we focus attention on the cooperative efforts that drive systems biology forward, highlighting the field's importance in answering basic biological questions, advancing medical research, and addressing societal issues.

Computational biology and mathematics work together to reveal the complexities of biological processes through mathematical modeling and the analysis of large biological datasets. This collaboration facilitates the understanding of biological phenomena at a systemic level and advances the field of systems biology. Physics is essential to the study of biological systems because it integrates concepts like thermodynamics, kinetics, and statistical mechanics to provide insights into molecular and cellular dynamics. Chemistry clarifies molecular interactions within biological systems by combining methods like proteomics, genome analysis, and metabolomics with data analysis in systems biology studies. Engineering principles and control theory provide engineering methodologies to investigate biological processes. Bioinformatics and data science are crucial for interpreting biological data using machine learning, network analysis, and data mining approaches. Data integration and analysis are made simple by tools and databases. Ecology and systems biology interact to provide insights into ecosystem dynamics, biodiversity, and the consequences of these discoveries for conservation biology and ecosystem management. Together, these fields of study enable us to comprehend the complex inner workings of ecological systems. Systems biology provides insights into illness causes and treatment development to enable customized medicine and combine clinical and molecular data to enhance patient care. Philosophical considerations on biological complexity and systems thinking ultimately inform ethical problems in systems biology research and applications, helping us better comprehend the nature of biological systems and life itself.

REFERENCES:

1. Hyduke, D. R., Lewis, N. E., & Palsson, B. Ø. (2013). Analysis of omics data with genome-scale models of metabolism. *Molecular BioSystems*.
2. Kitano, H. (2002). *Systems Biology: A Brief science overview*.
3. Alon, U. (2006). *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman & Hall/CRC.
4. Ideker, T., Galitski, T., & Hood, L. (2001). A New Approach to Deciphering the Complexity of Biological Systems. *Science*, 292(5519), 2001-2006.

5. Wagner, A. (2011). The molecular origins of evolutionary innovations. *Trends in Genetics*, 27(10).
6. Barabási, A. L., & Oltvai, Z. N. (2004). Network biology: Understanding the cell's functional organization. *Nature Reviews Genetics*.
7. Ideker, T., Galitski, T., & Hood, L. (2001). A new approach to decoding life: systems biology. *Annual Review of Genomics and Human Genetics*, 2(1), 343-372.

4. MATLAB COURSES

Based on the matlab certificate there are 20 courses I have completed

- Reinforcement Learning Onramp
- Image Processing with MATLAB
- App Building Onramp
- MATLAB Fundamentals
- Solving Nonlinear Equations with MATLAB
- Machine Learning with MATLAB
- Statistics Onramp
- Signal Processing Onramp
- MATLAB Programming techniques
- MATLAB for data Processing and Visualization
- Simulink Onramp
- Wireless Communications Onramp
- MATLAB Onramp
- Image processing Onramp
- Deep Learning Onramp
- Machine Learning Onramp
- Deep Learning with MATLAB
- Computer vision Onramp
- Object-Oriented Programming Onramp
- Signal Processing with MATLAB

MACHINE LEARNING ONRAMP:

Machine Learning Onramp, is a self-paced tutorial that provides an interactive introduction to practical machine learning methods for classification problems using hands-on exercises with automated assessments and feedback. Topics covered in Machine Learning Onramp include an overview of machine learning and the classification workflow, importing and preprocessing data, engineering features, and building classification models.

WHAT I LEARNT ?

- Participants will gain a solid understanding of the fundamental concepts, techniques, and algorithms used in machine learning. This includes supervised and unsupervised learning methods, data preprocessing techniques, model evaluation, and more.
- With a comprehensive understanding of core concepts such as supervised and unsupervised learning, coupled with proficiency in data preprocessing techniques and model evaluation, individuals gain the confidence to tackle real-world data challenges.

APP BUILDING ONRAMP:

App Building Onramp in MATLAB is an introductory course that teaches you how to create interactive applications in MATLAB using the App Designer tool. The App Designer allows users to build graphical user interfaces (GUIs) for their MATLAB programs, making it easier to interact with code, visualize results, and create more user-friendly applications. It creates the visual layout of an app, using drag-and-drop components like buttons, sliders, drop-down menus, plots, and other interactive elements.

WHAT I LEARNT?

- Many App Building Onramp courses emphasize collaboration and version control using tools like Git. We will learn how to work effectively in a team environment and manage code changes using version control systems.
- The course serves as a stepping stone for individuals who wish to pursue further learning and specialization in specific areas of app development, such as iOS or Android development, game development, or mobile app security.

MATLAB FUNDAMENTALS:

MATLAB Fundamentals is a course designed to provide an in-depth understanding of MATLAB, a high-level programming environment widely used for technical computing, data analysis, algorithm development, and visualization. The course typically covers the core aspects of MATLAB, including its language structure, built-in functions, data manipulation capabilities, and graphical features.

WHAT I LEARNT ?

- Develop proficiency in data analysis and visualization, learning how to import, manipulate, analyze, and present data effectively using MATLAB's powerful tools and functions.
- Modularization and function creation, enabling them to write maintainable and optimized code.
- Depending on the course's depth, participants may also gain exposure to advanced topics such as specialized toolboxes and symbolic computation, expanding their knowledge and skill set.

MATLAB ONRAMP:

MATLAB Onramp is a free, self-paced, online course designed to introduce users to the basics of MATLAB. It provides a hands-on learning experience through interactive exercises, allowing users to quickly build foundational skills in MATLAB programming and data analysis. Explore different data structures in MATLAB, such as arrays, matrices, cell arrays, and structures. Understand how to create and manipulate these data structures.

WHAT I LEARNT ?

- By the end of the course, learners should have a solid understanding of the MATLAB environment, syntax, and basic programming concepts. They will be able to navigate MATLAB confidently and write simple scripts to perform tasks like data analysis, visualization, and basic mathematical computations.
- MATLAB Onramp covers basic data analysis techniques, including manipulation of arrays and matrices, statistical calculations, and visualization of data.
- Completing the MATLAB Onramp serves as a solid foundation for diving into more advanced MATLAB topics and applications. Learners are equipped with the necessary knowledge and skills to explore more complex concepts, such as machine learning, signal processing, and advanced numerical methods.

SIGNAL PROCESSING WITH MATLAB:

Signal processing with MATLAB refers to using MATLAB to analyze, process, and manipulate signals. Signals can be audio, image, video, or other forms of data that vary with time or space. MATLAB provides a robust set of tools and functions specifically designed for signal-processing tasks. Understanding the fundamentals of signal processing, such as signals in time and frequency domains, sampling, quantization, and digital representation of signals.

WHAT I LEARNT ?

- Understand how signals are represented in MATLAB, including time-domain and frequency-domain representations.
- Explore different ways to visualize signals, such as plotting time-domain signals, spectrograms, and frequency spectra.
- Understand frequency-domain concepts and apply techniques like the Fast Fourier Transform (FFT) to analyze signal frequency components.

OBJECT-ORIENTED PROGRAMMING ONRAMP:

Object-Oriented Programming Onramp is a course designed to introduce the concepts of object-oriented programming (OOP) using MATLAB. This self-paced, interactive course covers the fundamentals of OOP and demonstrates how to apply these concepts to solve programming problems in MATLAB. Understand the basic concepts of OOP, including objects, classes, methods, properties, and encapsulation. Learn how these concepts help organize and structure code.

WHAT I LEARNT ?

- The Object-Oriented Programming (OOP) course equips with a robust skill set and a deep understanding of software development principles.
- Upon completion, students emerge with the ability to design, implement, and maintain complex software systems using OOP concepts such as encapsulation, inheritance, and polymorphism. They gain proficiency in organizing code efficiently, enhancing reusability, and ensuring scalability.
- Developing problem-solving skills and learn to navigate real-world software development challenges by applying design patterns and leveraging abstraction

SIMULINK ONRAMP:

Simulink Onramp is an introductory course designed to teach the basics of Simulink, a graphical programming environment used for modeling, simulating, and analyzing dynamic systems. Simulink is widely used in engineering and scientific applications, particularly for system modeling and simulation in areas like control systems, signal processing, robotics, and more. The Simulink Onramp course provides a hands-on, interactive learning experience to help users understand the fundamentals of Simulink and how to build basic models.

WHAT I LEARNT ?

- Understand the Simulink workspace and how to create and manage models.
- Adjust parameters in real-time during simulations to explore different system behaviors.
- Manipulate signals with operations like gain, delay, and filtering.

REINFORCEMENT LEARNING ONRAMP:

Reinforcement Learning Onramp is a course designed to introduce the basics of reinforcement learning (RL) using MATLAB and Simulink. Reinforcement learning is a type of machine learning where agents learn to make decisions by interacting with an environment and receiving feedback in the form of rewards or penalties. This course provides a hands-on introduction to the key concepts and tools used in reinforcement learning, guiding learners through the process of building, training, and testing RL agents.

WHAT I LEARNT ?

- Acquiring a comprehensive understanding of the fundamental concepts, principles, and terminology of reinforcement learning, including Markov Decision Processes (MDPs), value functions, policies, and exploration/exploitation tradeoffs.
- Help us to develop problem-solving skills necessary to tackle real-world reinforcement learning problems across diverse domains such as robotics, gaming, finance, healthcare, and more. learnt how to formulate RL problems, choose appropriate algorithms, and evaluate their performance.
- The course will enhance learners' critical thinking abilities as they analyze and evaluate the performance of different RL algorithms in various scenarios.

DEEP LEARNING ONRAMP:

Deep Learning Onramp is an introductory course designed to teach the fundamentals of deep learning using MATLAB. It provides a hands-on approach to learning the basics of deep learning concepts and how to implement them in MATLAB with the Deep Learning Toolbox. The course is typically

self-paced and focuses on key elements like neural networks, model training, and practical applications.

WHAT I LEARNT ?

- Many deep learning courses involve hands-on projects where students apply what they've learned to solve real-world problems. Completing these projects
- helps reinforce concepts and build a portfolio of work to showcase to potential employers.
- Completing a Deep Learning Onramp course can serve as a solid foundation for further study or employment in the field of deep learning and artificial intelligence. It can provide a starting point for more advanced courses or projects, or even entry-level positions in industries where deep learning expertise is valued.

MATLAB PROGRAMMING TECHNIQUES:

MATLAB Programming Techniques refer to a set of best practices and advanced methods used to develop robust, efficient, and maintainable MATLAB code. These techniques go beyond basic scripting and aim to help programmers write clean, optimized, and scalable code for various applications. Break down your code into smaller, reusable functions and scripts. This modularity improves code readability, reduces redundancy, and facilitates debugging.

WHAT I LEARNT ?

- Gaining a solid understanding of the MATLAB programming language, its syntax, and its various features.
- This course help us to develop algorithms and implement them in MATLAB for various applications, such as signal processing, image processing, control systems, and more.
- Understanding the importance of documenting their code and results effectively, including writing clear comments and generating reports or presentations.

MATLAB FOR DATA PROCESSING AND VISUALIZATION:

MATLAB for data processing and visualization involves using MATLAB's extensive tools and functions to analyze, transform, and visualize data. This application of MATLAB is popular among researchers, engineers, and data scientists due to its versatility in handling a wide range of data types and its ability to create detailed visualizations. MATLAB allows you to import data from various

sources, such as CSV files, Excel spreadsheets, text files, databases, and even web services. You can also export processed data to different formats for sharing or further analysis

WHAT I LEARNT ?

- This help us to learn how to create insightful visualizations of their data using MATLAB's plotting functions, including 2D and 3D plots, histograms, scatter plots, and more. They will also learn how to customize the appearance of their plots to effectively communicate their findings.
- Many courses include hands-on projects or assignments where participants apply their MATLAB skills to analyze real-world datasets and create meaningful visualizations.
- How to perform basic and advanced data analysis tasks using MATLAB, such as statistical analysis, curve fitting, signal processing, image processing, and machine learning.

IMAGE PROCESSING ONRAMP:

Image Processing Onramp is an introductory course designed to teach the basics of image processing using MATLAB. This course provides a hands-on learning experience through interactive exercises and is suitable for beginners or those with limited experience in image processing. It focuses on key concepts and techniques for manipulating, analyzing, and visualizing images in MATLAB.

WHAT I LEARNT ?

- Completing the Image Processing Onramp course can have numerous outcomes, depending on your goals and aspirations. To gain a solid understanding of the fundamental concepts and techniques used in image processing, equipping you with valuable knowledge applicable across various industries, from healthcare to robotics.
- This course can serve as a stepping stone for further exploration and specialization in fields like computer vision, machine learning, and artificial intelligence.
- Additionally, mastering image processing techniques can enhance your problem-solving skills, allowing you to tackle real-world challenges more effectively.

WIRELESS COMMUNICATION ONRAMP:

In MATLAB, the term "wireless communication onramp" is not a standard or widely recognized term. However, in the context of wireless communication systems and MATLAB, an "onramp" might refer to a set of tools, functions, or resources provided by MATLAB to facilitate the development,

simulation, or analysis of wireless communication systems. MATLAB offers various toolboxes and functions for wireless communication system design, simulation, and analysis.

WHAT I LEARNT ?

- Understand basic concepts like modulation, coding, and signal propagation.
- Learn how to generate various types of waveforms used in wireless communication, such as OFDM and QPSK.
- Simulate different types of communication channels, including multipath and fading channels.

IMAGE PROCESSING WITH MATLAB:

Image processing with MATLAB involves the manipulation and analysis of digital images using MATLAB's powerful built-in functions and toolboxes. MATLAB provides a comprehensive environment for performing a wide range of image processing tasks, from basic operations like loading and displaying images to advanced techniques such as image enhancement, segmentation, feature extraction, and pattern recognition.

WHAT I LEARNT?

- Understanding how images are represented in MATLAB, including grayscale, RGB, and other color spaces.
- Applying techniques to improve image quality, such as contrast adjustment, histogram equalization, and filtering for noise reduction.
- Measuring properties of image regions, such as area, centroid, and perimeter, for analysis and object recognition.

SIGNAL PROCESSING ONRAMP:

Signal Processing Onramp is an interactive tutorial provided by MATLAB that serves as an introductory course to the fundamentals of signal processing. It's designed for beginners who want to learn about signal processing concepts and techniques using MATLAB. The Onramp starts with an overview of what signal processing is and its importance in various fields such as

telecommunications, audio processing, and biomedical engineering. It covers fundamental concepts such as signals (continuous and discrete), systems, and basic operations like convolution and correlation.

WHAT I LEARNT ?

- Participants gain a strong foundational understanding of signal processing concepts and techniques, including time-domain and frequency-domain analysis, filtering, convolution, and spectral analysis.
- Implementing signal processing algorithms and techniques using MATLAB's tools and functions. They learn how to manipulate and analyze signals effectively, perform filtering operations, and conduct spectral analysis.

STATISTICAL ONRAMP:

The Statistics Onramp is an interactive learning module provided by MATLAB that serves as a beginner-level course covering the fundamentals of statistics using MATLAB's tools and functions. It's designed for individuals who want to learn about statistical concepts and their practical application in data analysis. The Onramp begins with an introduction to basic statistical concepts, such as descriptive statistics (mean, median, mode), variability (standard deviation, variance), and probability distributions.

WHAT I LEARNT ?

- Participants gain a strong foundational understanding of statistical concepts, including descriptive statistics, probability distributions, hypothesis testing, and regression analysis.
- Practical skills in using MATLAB for statistical analysis, learning how to perform various statistical tests, calculations, and visualizations efficiently.
- Interpreting data using statistical methods, enabling them to draw meaningful conclusions and make informed decisions based on data insights.

SOLVING NON-LINEAR EQUATIONS WITH MATLAB:

Solving nonlinear equations with MATLAB involves utilizing a variety of numerical methods and optimization algorithms to find solutions to equations that cannot be solved analytically. MATLAB provides a rich set of built-in functions and toolboxes specifically designed for this purpose, making it a powerful platform for tackling nonlinear equations encountered in various fields of science, engineering, and mathematics.

WHAT I LEARNT ?

- Participants become proficient in using a variety of numerical methods and algorithms for solving nonlinear equations, such as iterative methods, optimization techniques, and root-finding algorithms.
- Advanced MATLAB programming skills, learning how to implement and apply numerical methods to solve nonlinear equations efficiently.
- Participants acquire practical experience in solving a diverse range of nonlinear equation problems encountered in various fields, such as engineering, science, finance, and economics.

MACHINE LEARNING WITH MATLAB:

Machine learning with MATLAB offers a comprehensive platform for developing, implementing, and deploying machine learning models across a wide range of applications. MATLAB provides an extensive set of built-in functions, toolboxes, and resources specifically tailored for machine learning tasks, making it an ideal environment for both beginners and experienced practitioners in the field. MATLAB provides resources for newcomers to learn about machine learning concepts, including supervised learning, unsupervised learning, and reinforcement learning. Beginners can grasp fundamental concepts such as classification, regression, clustering, and dimensionality reduction through interactive tutorials, examples, and documentation.

WHAT I LEARNT ?

- Help us to Gain proficiency in utilizing MATLAB, a versatile computational tool, to develop, train, and deploy machine learning models for various realworld applications.
- With hands-on experience in data preprocessing, feature selection, model evaluation, and algorithm implementation, participants are equipped to tackle complex problems and make informed decisions in data-driven environments.
- Furthermore, acquiring the ability to interpret and communicate the results of machine learning models effectively, empowering them to contribute meaningfully to projects and initiatives across diverse industries.

DEEP LEARNING WITH MATLAB:

Deep learning with MATLAB provides a powerful and comprehensive environment for developing, implementing, and deploying deep neural network models for a wide range of applications. MATLAB offers an extensive set of built-in functions, toolboxes, and resources specifically tailored for deep learning tasks, making it an ideal platform for both beginners and experienced practitioners in the

field. MATLAB offers resources for newcomers to learn about deep learning concepts, including neural networks, convolutional neural networks (CNNs), recurrent neural networks (RNNs), generative adversarial networks (GANs), and deep reinforcement learning. Beginners can grasp fundamental concepts such as network architectures, training algorithms, and optimization techniques through interactive tutorials, examples, and documentation.

WHAT I LEARNT ?

- By harnessing deep learning algorithms within MATLAB's framework, researchers and practitioners have achieved remarkable outcomes. One notable area of impact is in image recognition and classification, where convolutional neural network (CNNs) have demonstrated unprecedented accuracy in identifying objects within complex visual scenes
- Deep learning with MATLAB has revolutionized medical image analysis, facilitating the detection of anomalies in MRI scans, X-rays, and histopathology slides. Researchers can leverage MATLAB's deep learning tools to create customized architectures tailored to specific medical imaging tasks, ultimately enhancing diagnostic accuracy and patient care.

COMPUTER VISION ONRAMP:

The Computer Vision Onramp is an interactive tutorial provided by MATLAB, offering an entrylevel course on the fundamental concepts and techniques of computer vision. Designed for beginners, this Onramp equips learners with the foundational knowledge needed to understand and apply computer vision algorithms using MATLAB's tools and functions. The Onramp begins with an overview of computer vision, explaining its significance, applications, and key concepts such as image representation, feature extraction, and object detection.

WHAT I LEARNT ?

- Upon completing the Computer Vision Onramp course, participants emerge with a robust understanding of the fundamental principles and techniques underlying computer vision.
- Armed with this knowledge, we gain the ability to process and analyze images effectively, extracting meaningful information and insights.
- To become proficient in employing various image processing techniques, feature extraction methods, and object detection algorithms to tackle real-world challenges.

