

PROFESSIONAL TRAINING REPORT

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

By

VIVEKAA A (43739038)



**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)

CATEGORY-I UNIVERSITY BY UGC

Accredited "A++" Grade by NAAC | 12B Status by UGC | Approved by AICTE

www.sathyabama.ac.in

DEPARTMENT OF BIOINFORMATICS AND DATA SCIENCE

BONAFIDE CERTIFICATE

This is to certify that the professional training report is the bonafide work of Vivekaa A (43739038) Who underwent professional training under the supervision of Dr. Swetha Sunkar from OCT 2023 to APR 2024.

Dr. SWETHA SUNKAR

INTERNAL GUIDE

Dr. H. JEMMY CHIRSTY

HEAD OF THE DEPARTMENT

Submitted for viva Examination held on:

INTERNAL EXAMINER

EXTERNAL EXAMINER

DECLARATION

I am **Vivekaa A** hereby declare that the Project Report entitled "PROFESSIONAL TRAINING REPORT" done by me under the guidance of **Dr.Swetha Sunkar** (Internal) and **Dr. Jemmy Chirsty** (External) at Sathyabama Institute of Science and Technology, Jeppiar Nagar, Rajiv Gandhi Road, Chennai-600119, Tamil Nadu, India, is submitted in partial fulfillment of the requirements for the award of **Master of Science degree in Bioinformatics and Data Science.**

DATE:

PLACE: CHENNAI

SIGNATURE OF THE CANDIDATE

ACKNOWLEDGEMENT

I am pleased to acknowledge my sincere thanks to **Board of Management of SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY** for their kind encouragement and support throughout the duration of my project. Their guidance has been invaluable in ensuring the successful completion of this project and I am sincerely grateful for their contribution.

I convey my thanks to **Dr. Rekha Chakaravarthy**, Dean, School of Science and Humanities for providing me necessary support and details at the right time during the progressive reviews.

I sincerely thank **Dr. H. Jemmy Chirsty**, Head of the department, Dept. of Bioinformatics, for her unwavering support and timely provision of necessary details during the progressive reviews. Her support has been instrumental in helping me to overcome obstacles and achieve success in this project.

I am deeply indebted to my project guide **Dr. Swetha Sunkar**, for her invaluable guidance, suggestions and constant encouragement throughout the project. Her expertise and insights have been instrumental in helping my work and ensuring its successful completion.

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

VIVEKAA A

TABLE OF CONTENTS

S.NO	CONTENTS	PAGE NO.
1	DEEP LEARNING INTERNSHIP (VALUE ADDED COURSE)	1
1.1	ABSTRACT	1
1.2	INTRODUCTION	2
1.3	OBJECTIVES	3
1.4	INTERNSHIP ACTIVITIES	4
1.5	APPLICATION OF DEEP LEARNING	5
1.6	KEY CONCEPTS LEARNED	6
1.7	SUMMARY	11
2	EXPLORING GENETIC VARIATION IN NOTUM AND THEIR IMPACT ON VALSARTAN BINDING – MINI PROJECT	12
2.1	ABSTRACT	12
2.2	INTRODUCTION	13
2.3	AIM AND SCOPE	14
2.4	MATERIALS AND METHODS	15
2.5	RESULT AND DISCUSSION	16
2.6	CONCLUSION	21
2.7	REFERENCES	22
3	MATLAB SELF PACED ONLINE COURSES	24
4	BIOINFOBYTE - E-NEWSLETTER	33
5	BOOK CHAPTER INTERDISCIPLINARY PERSPECTIVES IN SYSTEMS BIOLOGY:FROM THEORY TO PRACTICE - ISBN Number: 978-93-340-3987-0	36
6	NATIONAL CONFERENCE ON TRANSFORMING HEALTHCARE: ADVANCES IN BIOINFORMATICS AND BIOLOGICAL SCIENCES (ABBS - 2024)	42



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)

CATEGORY - 1 UNIVERSITY BY UGC



Accredited with "A++" Grade by NAAC | 12B Status by UGC | Approved by AICTE



CERTIFICATE

This certificate is presented to Vivekaa.A from MSc.Bioinformatics and DataScience,1st_year for successfully completing the three months Value Added Course (Internship) on Deep Learning held from September 2023 to November 2023 , organized by the Dept. of Computer Science & Engineering, Sathyabama Institute of Science & Technology.

Dr.T.Sasikala

Dean

(School of computing)

Dr. S. Vigneshwari

HOD

(School of computing)

Mrs.Parveen.A

(Faculty Coordinator)

Mrs.R.Sudha

(Faculty Coordinator)

1. DEEP LEARNING INTERNSHIP (VALUE ADDED COURSE)

CHAPTER 1

1.1 ABSTRACT

Deep learning is a subset of machine learning that focuses on training neural networks to learn and make decisions from vast amounts of data. This internship report provides a comprehensive overview of my experiences and learnings during an internship in the field of deep learning. The report details the objectives of the internship, activities undertaken, key learnings, and insights gained through hands-on experience with deep learning frameworks and applications. The report emphasizes the practical applications of deep learning technology across various domains and data analysis.

Deep learning has emerged as a powerful approach to solving complex problems in artificial intelligence (AI) by leveraging deep neural networks to learn hierarchical representations from data. The internship in deep learning aimed to provide practical exposure to cutting-edge techniques and tools used in this field. The primary objectives were to gain proficiency in deep learning frameworks, understand neural network architectures, and apply deep learning techniques to real-world problems. This report outlines the activities, projects, and outcomes achieved during the internship period.

CHAPTER 2

1.2 INTRODUCTION

During my internship in the field of deep learning, I had the incredible opportunity to learn evolving area of artificial intelligence (AI) that is revolutionizing various industries, from healthcare and finance to autonomous vehicles and natural language processing. Deep learning, a subset of machine learning, focuses on developing and training neural networks to learn and make decisions from large volumes of data. Throughout this experience, I encountered a diverse range of applications. For instance, in healthcare, deep learning models are enhancing diagnostics through image analysis and predictive analytics, revolutionizing patient care by enabling earlier disease detection and personalized treatment plans. Finance is another domain benefiting from deep learning's capabilities utilizing predictive algorithms to optimize investment strategies and detect fraud patterns with greater accuracy.

Reflecting on my internship, I gained valuable insights into the intricacies of deep learning technologies. I learned about the underlying principles of neural networks and the training process, understanding how these algorithms extract patterns from data to make informed predictions. I also encountered challenges in implementing deep learning solutions, such as data quality issues and computational complexity. Moreover, I became more aware of the ethical considerations in AI, emphasizing the need for responsible deployment to mitigate biases and ensure transparency.

CHAPTER- 3

1.3 OBJECTIVES:

The primary objectives of my internship were:

1. Gain hands-on experience with deep learning frameworks and tools.
2. Understand the theoretical foundations of neural networks and their applications.
3. Apply deep learning techniques to solve real-world problems.
4. Collaborate with professionals and contribute to ongoing projects in the field.

CHAPTER 4

1.4 INTERNSHIP ACTIVITIES

1. Learning Deep Learning Fundamentals
2. Studied neural network architectures, including feedforward networks, convolutional neural networks (CNNs), recurrent neural networks (RNNs), and generative adversarial networks (GANs).
3. Explored optimization algorithms such as stochastic gradient descent (SGD), Adam, and RMSprop for training deep neural networks.
4. Understood activation functions, loss functions, and regularization techniques used in deep learning models.

CHAPTER- 5

1.5 APPLICATION OF DEEP LEARNING

1. Collaborated with the team to develop and deploy deep learning solutions for real-world problems.
2. Contributed to projects involving computer vision tasks like object detection, image segmentation, and facial recognition.
3. Explored natural language processing (NLP) applications such as sentiment analysis, text generation, and language translation using deep learning models.

CHAPTER-6

1.6 KEY CONCEPTS LEARNED:

1.6.1 BASICS OF DEEP LEARNING:

Deep learning, a subset of machine learning, enables computers to accomplish tasks by learning from examples, akin to human learning. To illustrate, if we want to teach a computer to recognize cats, we don't specify features like whiskers or tails instead, we expose it to numerous cat images. Through this exposure, the computer autonomously discerns common patterns and learns to identify cats. This methodology hinges on neural networks, inspired by the human brain, composed of layers of interconnected nodes that process data. Greater network depth enables learning of increasingly intricate features, empowering the system to tackle more complex tasks.

1.6.2 ADVANCING MACHINE LEARNING-DEEP LEARNING

Machine learning, a subset of artificial intelligence, has reshaped computing by enabling systems to learn from data and make decisions independently. Traditionally, machine learning techniques required manual feature engineering, where domain experts meticulously selected and transformed variables for model optimization.

However, deep learning, a specialized branch of machine learning, has revolutionized this process by employing neural networks with multiple layers to automatically extract relevant features from raw data.

Deep learning differs from traditional machine learning in its utilization of deep neural networks, inspired by the human brain's structure. These networks autonomously learn from vast amounts of data, reducing the need for manual intervention in feature engineering. This advancement accelerates model development and deployment, paving the way for handling complex tasks with greater efficiency and accuracy.

While feature engineering remains crucial in both traditional machine learning and deep learning methodologies, the transition to deep learning signifies a paradigm shift in artificial intelligence. By automating feature extraction and learning from raw

data, deep learning algorithms unlock unprecedented potential in various domains, including image recognition, natural language processing, and autonomous systems. As research in deep learning progresses, its impact on society and technology is poised to reach new heights, shaping the future of AI-driven innovation.

1.6.3 IMPORTANCE OF DEEP LEARNING

Handling Unstructured Data: Deep learning excels at learning from unstructured data, such as images, text, and audio, without the need for extensive preprocessing to standardize the data.

Processing Large Data Sets: With the advent of GPUs, deep learning models can efficiently process vast amounts of data at high speeds, enabling the analysis of big datasets that were previously challenging to handle.

High Accuracy: Deep learning models consistently deliver highly accurate results in complex tasks like computer vision, natural language processing (NLP), and audio analysis, surpassing traditional machine learning methods.

Automated Pattern Recognition: Deep learning models have the unique ability to automatically detect and learn complex patterns within data, reducing the need for manual feature engineering by machine learning engineers.

1.6.4 CORE CONCEPTS OF DEEP LEARNING:

Foundation of deep learning is Neural networks are computational models inspired by the structure and functioning of the human brain. They are comprised of interconnected nodes, or "neurons," which mimic the behavior of biological neurons. These neurons work collaboratively to process information and generate outputs based on the input data. Just like different regions of the human brain specialize in various tasks, neural networks are organized into layers, with each layer performing specific functions in the overall computation process.

What sets deep neural networks apart is their depth, which refers to the number of layers between the input and output layers. These networks have multiple hidden layers that allow them to learn complex patterns and features from the input data. By incorporating more layers, deep neural networks can capture intricate

relationships within the data, leading to more accurate predictions and outcomes. This depth is a key factor in the success of deep learning algorithms in handling complex tasks such as image recognition, natural language processing, and speech recognition.

Activation functions are essential components of neural networks that introduce non-linearity into the model. They determine the output of each neuron based on the weighted sum of its inputs and bias. By applying activation functions to the outputs of neurons, neural networks can learn complex mappings between the input and output data. These functions introduce non-linearities into the network, allowing it to model and understand complex relationships within the data. Common activation functions include sigmoid, tanh, ReLU (Rectified Linear Unit), and softmax, each with its own characteristics and suitability for different types of tasks.

1.6.5 HOW DEEP LEARNING WORKS

Deep learning operates by using neural networks to progressively extract and classify features from data, allowing it to make accurate predictions or classifications. Here's a breakdown of how it works:

Feature Extraction: Deep learning models automatically extract relevant features from the raw input data. For example, in a cats and dogs classification task, the model might identify important features like the eyes, face, and body shape of the animals.

Neural Network Architecture: Deep learning models are constructed using deep neural networks, which consist of multiple layers. A basic neural network comprises an input layer, one or more hidden layers, and an output layer. Deep learning models have multiple hidden layers, hence the term "deep," which allows them to learn complex patterns and representations.

Processing through Layers: The input layer receives the raw data, which is then passed through the nodes (neurons) of the hidden layers. Each layer processes the data, transforming it into increasingly abstract representations. As the data passes through more layers, the model can discern more intricate patterns and relationships.

Hierarchical Representation: Deep learning models create a hierarchical representation of the input data. The initial layers capture basic features, while subsequent layers combine these features to form more sophisticated representations that align closely with the desired output.

Output Layer: The output layer utilizes the information processed by the hidden layers to make predictions or classifications. For instance, in an image classification task, the output layer might determine the likelihood that the input image depicts a dog or a cat.

By leveraging this hierarchical feature extraction process within deep neural networks, deep learning models can achieve high accuracy in tasks ranging from image and speech recognition to natural language understanding and decision-making. This ability to automatically learn intricate patterns and representations from data is what makes deep learning so powerful and widely applicable across various domains.

1.6.6 APPLICATIONS:

Deep learning has become integral to a wide array of technology applications, powering various artificial intelligence solutions that are reshaping industries across the globe. These applications demonstrate the remarkable potential of deep neural networks in solving complex problems efficiently.

One prominent area where deep learning excels is computer vision (CV). In self-driving cars, for instance, computer vision enables the detection of objects on the road, aiding in collision avoidance. Additionally, CV is utilized in tasks like face recognition, pose estimation, image classification, and anomaly detection, enhancing security and automation systems.

1. Automatic Speech Recognition (ASR) is another crucial domain benefiting from deep learning. ASR technologies, widely used in smartphones and virtual assistants like Google and Siri, enable seamless voice commands, text-to-speech conversion, audio classification, and voice activity detection.
2. Generative AI has witnessed a surge in demand, exemplified by the success of CryptoPunk NFTs, a generative art collection created using deep learning

models. Tools have revolutionized text generation, enabling novel writing and code generation for data science projects.

3. Deep learning is also transforming translation technologies, enabling the translation of not only languages but also photos to text using Optical Character Recognition (OCR) and text to images using NVIDIA's GauGAN2.
4. In the financial sector, time series forecasting powered by deep learning is essential for predicting market crashes, stock prices, and weather changes, providing valuable insights for decision-making.
5. Automation is a major application of deep learning, seen in tasks such as training robots for warehouse management and even achieving groundbreaking milestones like OpenAI's Dota AI beating professional gaming teams.
6. For customer service applications, deep learning is utilized in handling customer feedback and providing seamless interactions through chatbot applications.
7. Finally, the field of biomedicine has seen significant advancements with deep learning, facilitating cancer detection, drug development, anomaly detection in medical imaging like chest X-rays, and enhancing medical equipment performance.

These applications underscore the versatility and impact of deep learning across diverse sectors, highlighting its role in driving innovation and efficiency in modern technological solutions.

CHAPTER - 7

1.7 SUMMARY

Deep learning represents a transformative approach within the broader field of artificial intelligence and machine learning. By leveraging neural networks with multiple layers, deep learning algorithms autonomously extract complex features from raw data, revolutionizing tasks that were once reliant on manual feature engineering. This paradigm shift has accelerated the development and deployment of advanced AI models across various domains, from computer vision and natural language processing to robotics and biomedical applications.

The importance of deep learning lies in its ability to handle unstructured data efficiently, process vast datasets with GPUs, and consistently deliver high accuracy in complex tasks. Deep learning frameworks like Keras have democratized access to these technologies, enabling researchers and developers to experiment and innovate rapidly.

As deep learning research progresses, its impact on society continues to expand, paving the way for new AI-driven solutions that enhance automation, decision-making, and creativity. With ongoing advancements in neural network architectures and learning techniques, the future of deep learning promises to unlock even greater capabilities, driving innovation and reshaping the landscape of intelligent systems.

EXPLORING GENETIC VARIATION IN NOTUM AND THEIR IMPACT ON VALSARTAN BINDING – MINI PROJECT

CHAPTER-1

2.1 ABSTRACT

This study explores the impact of a single nucleotide polymorphism (SNP) within the NOTUM gene, a key regulator of the Wnt signaling pathway and a potential therapeutic target, on the binding affinity of valsartan, an angiotensin II receptor antagonist used in hypertension and heart failure. The central question addressed is whether specific NOTUM SNPs alter the binding affinity of valsartan to its target receptors. While NOTUM's role in Wnt signaling and therapeutic potential is established, the influence of NOTUM genetic variations on drug response, particularly with valsartan, remains underexplored.

Motivated by the need to understand genetic determinants of drug response, this research uses computational docking studies to compare the binding interactions of valsartan with native and mutant NOTUM models. The study aims to bridge a critical gap in understanding the genetic basis of drug efficacy without in vitro assays. Preliminary findings suggest that certain NOTUM SNPs may influence the affinity of valsartan for its target receptors, potentially contributing to variability in drug response among individuals.

These findings emphasize the importance of integrating genetic information into drug development and personalized medicine approaches. Understanding how specific NOTUM SNPs affect valsartan affinity can inform pharmacogenomic strategies aimed at optimizing drug dosing and treatment outcomes, ultimately advancing precision medicine and minimizing adverse drug reactions in hypertensive patients. This research underscores the significance of leveraging genetic insights to enhance therapeutic efficacy and patient care in cardiovascular medicine.

CHAPTER-2

2.2 INTRODUCTION:

In the field of drug development and pharmacogenomics, understanding genetic variations such as single nucleotide polymorphisms (SNPs) within drug target genes is essential for assessing drug responses and optimizing therapeutic outcomes. SNPs are common genetic changes involving single nucleotide base substitutions that can affect protein structure and function, impacting drug interactions within the body.

When SNPs occur within genes encoding drug targets, they can significantly alter drug binding affinity and influence treatment efficacy. The location of SNPs within critical regions of drug target genes may lead to amino acid substitutions, changes in protein shape, or disruptions at active binding sites. As a result, different genetic variants of drug targets can exhibit varying responses to medications, affecting treatment effectiveness and dosage requirements.

Pharmacogenomic studies focusing on SNP localization within drug target genes provide insights into the molecular mechanisms underlying individual variations in drug response. By identifying specific SNP variants and their impact on drug-target interactions, researchers can optimize drug design and dosing strategies. This approach enhances therapeutic efficacy and helps minimize adverse drug reactions.

This study aims to investigate the localization and functional consequences of SNPs within a proposed drug target, evaluating their influence on drug binding affinity and pharmacological responses. Understanding the genetic basis of drug - target interactions can lead to improved drug development strategies and more tailored therapeutic approaches. These insights contribute to optimizing clinical practice and advancing drug discovery efforts.

CHAPTER – 3

2.3 AIM AND SCOPE

Investigating the potential impact of a specific genetic variation within the NOTUM gene on the binding affinity of valsartan, a medication used to treat hypertension and heart failure. By utilizing computational docking studies, the research aims to compare how valsartan interacts with both the normal and mutant forms of the NOTUM protein. This approach allows for an exploration of how genetic differences in NOTUM may influence the effectiveness of valsartan in binding to its target receptors.

The study seeks to contribute insights into the molecular mechanisms underlying drug response variability, particularly concerning valsartan therapy. Understanding how specific genetic variations within NOTUM affect valsartan binding could have implications for personalized medicine strategies, potentially informing optimized treatment approaches and minimizing the risk of adverse drug reactions in patients with hypertension.

Ultimately, the research aims to advance knowledge in pharmacogenomics, highlighting the importance of considering genetic factors in drug development and therapeutic decision-making processes, particularly in cardiovascular medicine.

CHAPTER - 4

2.4 MATERIALS AND METHODS:

The structure of NOTUM - VALSARTAN complex was studied and downloaded from RCSB PDB (<https://www.rcsb.org/>). List of possible variants recorded for notum is collected from ENSEMBL (<https://www.ensembl.org/index.html>) from the listed variants 5 SNPs were chosen based on its impact and position in the protein ASP88ASN, ASN176LYS, GLY235ALA THR236HIS, LEU429SER. The structure of VALSARTAN (drug) was downloaded from PubChem (<https://pubchem.ncbi.nlm.nih.gov/>).

The drug - target complex was loaded in discovery studio and the interacting sites of notum with valsartan was mapped and the mutant model was built using BUILT MUTANT tool for each possible SNPs.

The mutant model was docked with the valsartan in the same site where the drug interacts with native protein using CDOCKER method and the interacting site and pattern was studied for each mutant model.

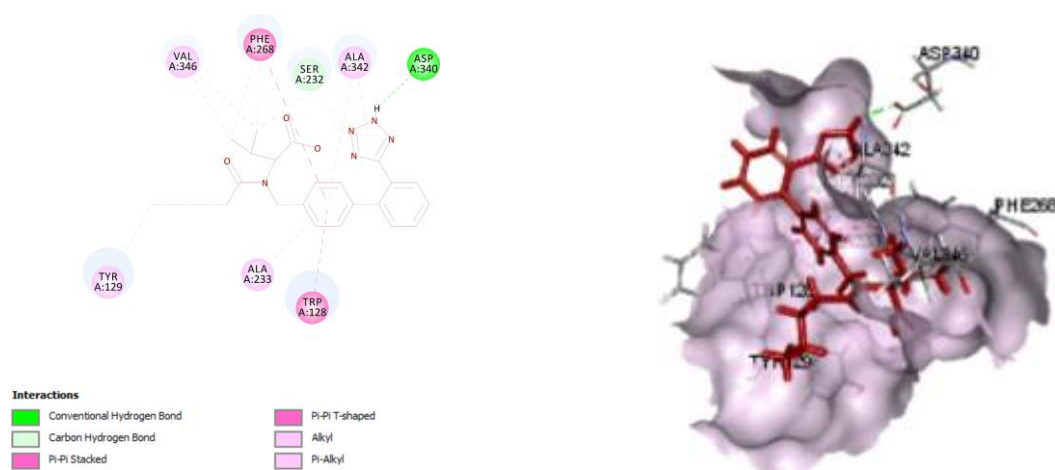
2.5 RESULT AND DISCUSSION:

2.5.1 ASP TO ASN at 88th position:

Aspartic acid (D) is a negatively charged amino acid with a carboxyl group in its side chain, which can form hydrogen bonds and participate in electrostatic interactions with other molecules. In contrast, asparagine (N) is an uncharged polar amino acid with an amide group in its side chain, capable of forming hydrogen bonds but lacking the negative charge of aspartic acid.

The substitution of aspartic acid (D) by asparagine (N) at position 88 in NOTUM can lead to structural and chemical changes at the binding site. This mutation can alter the local environment's polarity, hydrogen bonding potential, and electrostatic interactions. Consequently, the interaction between VALSARTAN and the mutated NOTUM protein may be affected, potentially impacting the binding affinity and stability of the drug-protein complex.

Furthermore, changes in amino acid properties due to the ASP88ASN mutation can influence the conformational flexibility and dynamics of the protein-drug interaction, ultimately influencing the functional consequences of the genetic variation on drug response and efficacy.



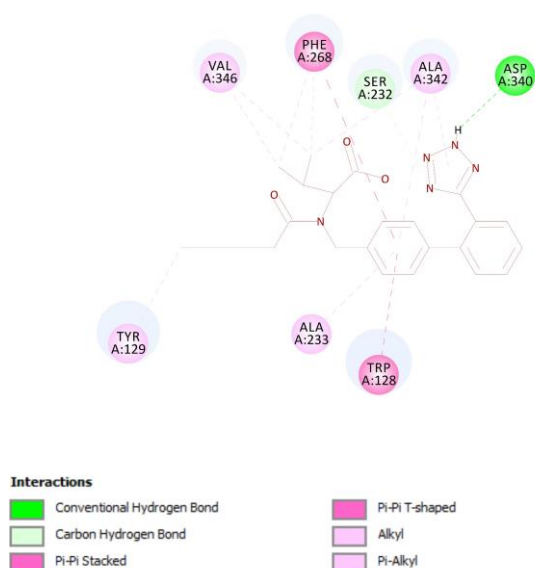
The interaction of drug with this mutant protein results in disruption of interaction between the drug and few residues in the protein, they are val,pro,ile,phe,glu,thr.All these residues promote drug binding affinity,because of this mutation the drug

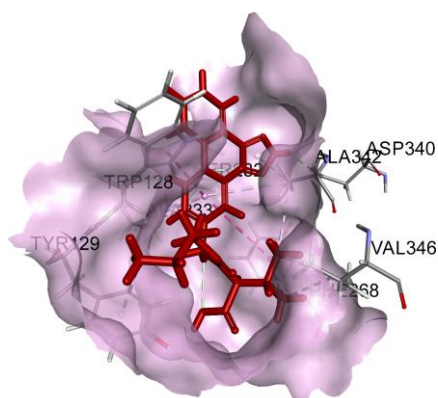
target complex becomes less stable. Protein structure, function and drug binding affinity alters.

2.5.2 ASP TO LYS AT 176th POSITION

The substitution of asparagine (N) with lysine (K) at position 176 in the NOTUM protein can have profound implications for drug-protein interactions, particularly with VALSARTAN, due to the distinct chemical properties of these amino acids.

Asparagine is a polar, neutral amino acid with an amide group in its side chain, facilitating hydrogen bonding interactions. In contrast, lysine is a positively charged amino acid at physiological pH, possessing a larger, highly polar side chain capable of strong electrostatic interactions with negatively charged groups. This mutation can lead to changes in the local conformation and charge distribution within the drug-binding site, potentially altering the spatial arrangement and accessibility of nearby residues and ligands like VALSARTAN. The introduction of lysine's positive charge may disrupt existing hydrogen bonds or introduce new electrostatic interactions, affecting the binding affinity and stability of the drug-protein complex. Consequently, altered interactions resulting from the ASN176LYS mutation could impact the pharmacological activity, metabolism, or overall therapeutic efficacy of VALSARTAN.

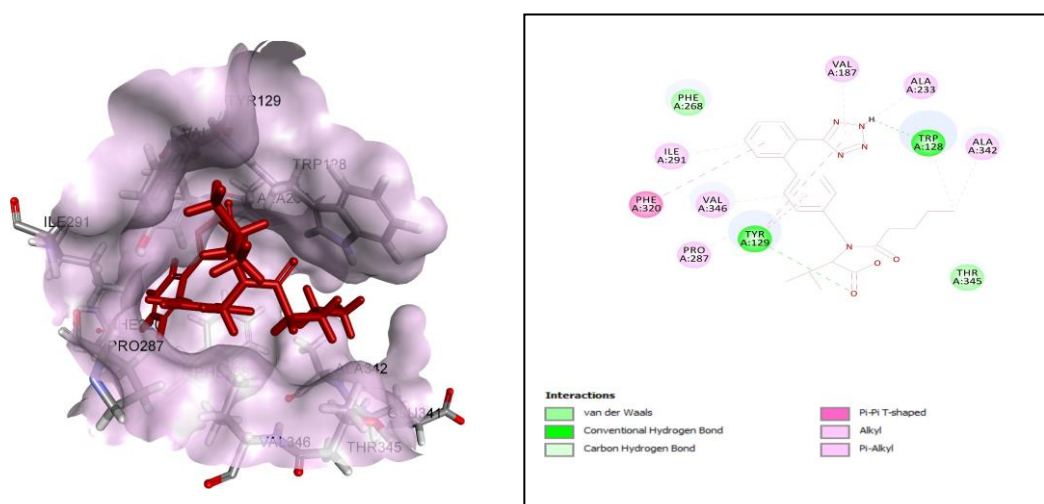




Due to this variation interaction at two sites are getting disrupted. The drug does not interact with Val and Thr, which increases the stability of the interaction, hence the stability of the interaction decreases.

2.5.3 GLY TO ALA AT 235th POSITION:

The substitution of glycine (G) with alanine (A) at the 235th position in the NOTUM protein, known as the GLY235ALA mutation, can impact drug-target interactions due to the distinct properties of these amino acids. Glycine is the smallest amino acid among the standard twenty amino acids, consisting of just a hydrogen atom as its side chain, which makes it highly compact and flexible. In contrast, alanine is more rigid because of its larger methyl group side chain. This substitution can alter the local protein conformation, potentially affecting the shape and dynamics of the drug-binding site within NOTUM.

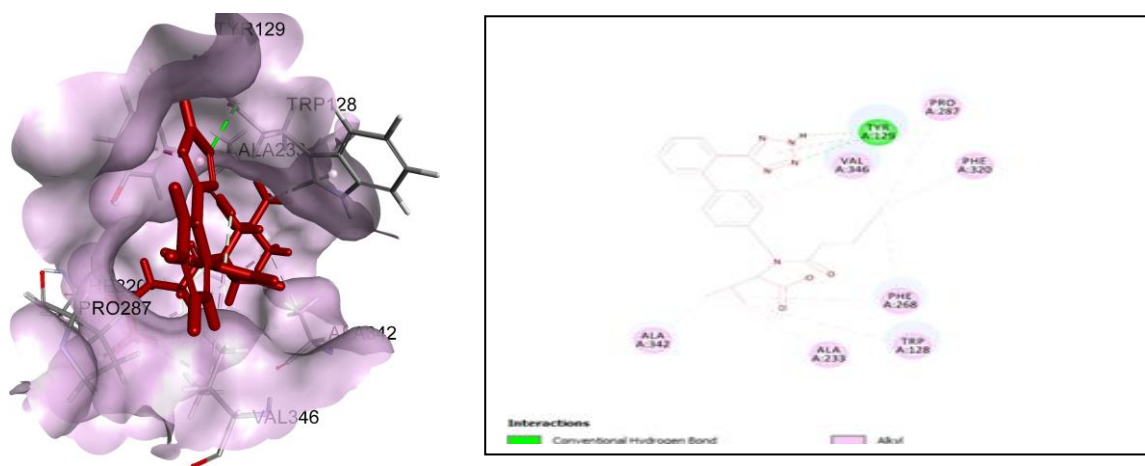


The interaction between drug and protein compared to the mutant model is similar as there is only one difference in the mutant model interaction. The mutant model

does not interact with Glu at 341st site which plays a significant role in drug efficacy and specificity.

2.5.4 THR TO HIS AT 236th POSITION:

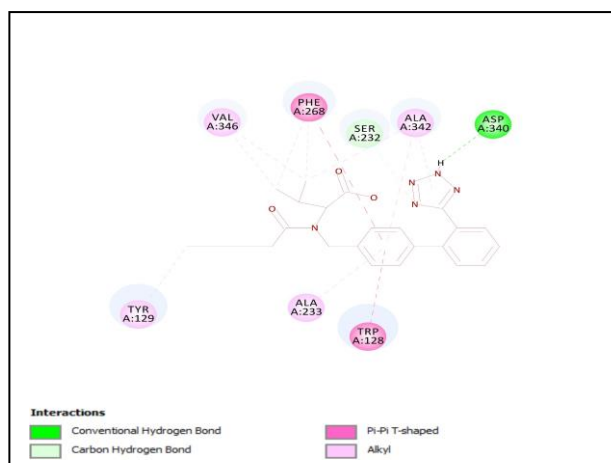
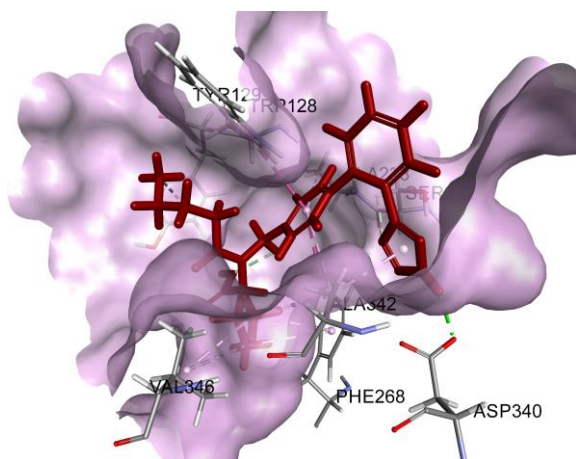
Threonine, being polar with a hydroxyl group (-OH) in its side chain, typically forms hydrogen bonds crucial for stabilizing protein-ligand complexes. In contrast, histidine is positively charged at physiological pH and features an imidazole ring capable of participating in diverse interactions like cation- π bonds and alternative hydrogen bonding patterns. This substitution may alter the binding affinity and mode of interaction between VALSARTAN and the mutant NOTUM protein, potentially affecting the local protein conformation and dynamics around the drug-binding site.



In mutant model drug interaction at Val,Ile,Glu,Thr which promotes drug induced modulation of protein function hence the stability of interaction gets reduced.

2.5.5 LEU TO SER AT 429th POSITION:

Leucine is hydrophobic and typically involved in stabilizing protein structures and interactions within hydrophobic cores. In contrast, serine introduces polarity with its hydroxyl group, potentially altering the local environment by forming new hydrogen bonds or disrupting existing hydrophobic interactions. This mutation can affect the shape, size, and chemical properties of the binding pocket, consequently influencing the binding affinity and specificity of drugs like VALSARTAN for the mutated NOTUM protein.



In mutant model , The mutation of Ser from Leu affects the interaction of drug with Val,Pro,Ile,Phe,Glu,Thr,which makes the complex less stable.

CHAPTER - 6

2.6 CONCLUSION:

This study employed computational modeling techniques to investigate the interaction between the enzyme NOTUM and the drug VALSARTAN, with a focus on understanding the impact of genetic variants (SNPs) within the NOTUM protein. The crystal structure of the NOTUM-VALSARTAN complex was obtained from the RCSB Protein Data Bank (PDB), and potential NOTUM variants were sourced from ENSEMBL, leading to the selection of five key SNPs (ASP88ASN, ASN176LYS, GLY235ALA, THR236HIS, LEU429SER) for further analysis. Molecular modeling was conducted using Discovery Studio software to generate mutant models of NOTUM corresponding to each SNP, and subsequent docking studies with VALSARTAN were performed to assess alterations in binding affinity and interaction patterns compared to the wild-type protein. These computational findings provide important insights into potential variations in drug response associated with genetic variability in NOTUM, underscoring the significance of considering genetic factors in pharmacological research and personalized medicine applications. Further experimental validation of these computational predictions would be essential for translating these findings into clinical practice.

CHAPTER - 7

2.7 REFERENCES:

1. Guo, F., Seldin, M., Péterfy, M. et al. NOTUM promotes thermogenic capacity and protects against diet-induced obesity in male mice. *Sci Rep* 11, 16409 (2021). <https://doi.org/10.1038/s41598-021-95720-1>
2. Zhao Y, Jolly S, Benvegnu S, Jones EY, Fish PV. Small-molecule inhibitors of carboxylesterase Notum. *Future Med Chem.* 2021 Jun;13(11):1001-1015. doi: 10.4155/fmc-2021-0036. Epub 2021 Apr 22. PMID: 33882714; PMCID: PMC8130783.
3. Vaduganathan M, Mentz RJ, Claggett BL, Miao ZM, Kulac IJ, Ward JH, Hernandez AF, Morrow DA, Starling RC, Velazquez EJ, Williamson KM, Desai AS, Zieroth S, Lefkowitz M, McMurray JJV, Braunwald E, Solomon SD. Sacubitril/valsartan in heart failure with mildly reduced or preserved ejection fraction: a pre-specified participant-level pooled analysis of PARAGLIDE-HF and PARAGON-HF. *Eur Heart J.* 2023 Aug 14;44(31):2982-2993. doi: 10.1093/eurheartj/ehad344. PMID: 37210743; PMCID: PMC10424880.
4. Pontremoli R, Borghi C, Perrone Filardi P. Renal protection in chronic heart failure: focus on sacubitril/valsartan. *Eur Heart J Cardiovasc Pharmacother.* 2021 Sep 21;7(5):445-452. doi: 10.1093/ehjcvp/pvab030. PMID: 33822031; PMCID: PMC8453284.
5. Ross JS, Schenkein DP, Kashala O, Linette GP, Stec J, Symmans WF, Pusztai L, Hortobagyi GN. Pharmacogenomics. *Adv Anat Pathol.* 2004 Jul;11(4):211-20. doi: 10.1097/01.pap.0000131825.77317.ee. PMID: 15220824.
6. Bakar NS. Pharmacogenetics of common SNP affecting drug metabolizing enzymes: comparison of allele frequencies between European and Malaysian/Singaporean. *Drug Metab Pers Ther.* 2021 Mar 18. doi: 10.1515/dmdi-2020-0153. Epub ahead of print. PMID: 33735954.
7. Xu H, Kang Y, Liang T, Lu S, Xia X, Lu Z, Hu L, Guo L, Zhang L, Huang J, Ye L, Jiang P, Liu Y, Xinyi L, Zhai J, Wang Z, Liu Y. SNP-based and

- haplotype-based genome-wide association on drug dependence in Han Chinese. *BMC Genomics*. 2024 Mar 6;25(1):255. doi: 10.1186/s12864-024-10117-4. PMID: 38448893; PMCID: PMC10919046.
8. Petrovski S, Szoek CE, Sheffield LJ, D'souza W, Huggins RM, O'brien TJ. Multi-SNP pharmacogenomic classifier is superior to single-SNP models for predicting drug outcome in complex diseases. *Pharmacogenet Genomics*. 2009 Feb;19(2):147-52. doi: 10.1097/FPC.0b013e32831d1dfa. PMID: 19077679.

3. BIOINFOBYTE - E-NEWSLETTER:

In my article titled "Decoding the Human Blueprint: Advances in Genomic Research," I provided an in-depth exploration of the field of genomics and its transformative impact on medical science. The article commenced with an introduction to genomics, elucidating the study of genomes and their fundamental role in understanding genetic information at a molecular level. I discussed the significance of genomic variations in contributing to various diseases, highlighting the importance of studying these variations for personalized medicine and targeted therapies.

A central focus of the article was the Human Genome Project, a historic milestone in scientific research that revolutionized our understanding of human genetics. This ambitious endeavor to sequence the entire human genome laid the foundation for subsequent advancements in genomics and paved the way for breakthroughs in personalized medicine and genetic diagnostics.

Next, I delved into the advancements in next-generation sequencing technologies, which have significantly accelerated genome sequencing processes, making them more efficient and accessible. This technological progress has facilitated large-scale genomic studies and enabled researchers to unravel complex genetic interactions underlying disease susceptibility and treatment response. Furthermore, I explored the revolutionary CRISPR-Cas9 genome editing technology and its profound implications for genetic research and therapeutic applications. CRISPR-Cas9 has emerged as a powerful tool for precise gene editing, offering unprecedented opportunities to correct genetic defects and potentially cure genetic diseases.

Looking towards the future, I emphasized the promising prospects of genomic medicine, where tailored treatments and preventive strategies can be developed based on an individual's unique genetic makeup. Personalized medicine holds immense potential for improving patient outcomes by optimizing treatment approaches and minimizing adverse effects.

VIVEKAA.A I YEAR M.SC BIOINFORMATICS & DATA SCIENCE	DECODING THE HUMAN BLUEPRINT: ADVANCES IN GENOMIC RESEARCH	#ARTICLE 2
--	---	------------

Introduction to Genomics:

From Bacteria to blue whale all living organisms have their own genome. Each genome is unique and each contains the information needed to build and maintain organisms, throughout its life. Scientists have advanced from studying individual or a limited number of genes to exploring thousands of genes exclusively within the past few decades. This progression signifies a shift from analysing isolated genetic units to investigating the entire genome of an organism. The study of entire genetic material or genome "Genomics", initially focused to unravel the entire genetic sequence of different species which is called as sequencing (Determining the precise order of nucleotides within a DNA/RNA molecule), has expanded toward a more functional level (i.e.,) Studying the expression profiles and the roles of proteins and genes. Genomics has evolved from mapping individual genes and identifying genetic variations to genome engineering through CRISPR technology and other methods for accurate gene editing to achieve the milestone of developing personalised medicine for all diseases.

In 2003, the Human Genome Project produced a genome sequence that accounted for over 90% of the human genome. It was as close to complete as the technologies for sequencing DNA allowed at the time. Published draft involved 10 years of hard work by the six countries of IHGSC and 3 years or less work by Celera associates. The combined achievements has been hailed with applause all over the world as it unfold the "Mysteries of Life" which was hidden in a magic box. Later on many descriptions have been given for the achievement as "Supplying a Periodic table of life", "Providing Library of life" and finally "The Holy Grail of Human Genetics."

Genomic Variation and Disease:

It is undeniably evident that every individual is distinct, displaying variations in physical attributes such as complexion, eye colour, height, and voice. Remarkably, around 99.6% of the human genome is identical between unrelated individuals, with ~ 0.4% accounting for genetic variations that contribute to diversity.

Many genetic variations either result in subtle differences or merely contribute to our distinct appearances. Some genetic variations, however, have the potential to lead to diseases. There are different type of genetic variations which includes Single Base-Pair Substitution where a single base pair is substituted, it can be either transition or transversion, Insertion or deletion which is popularly called as indel and Structural Variations which includes Insertion, deletion, Inversion, copy number variations and Duplication. Variations in the human genome play a crucial role in medicine, with only a small portion of these variants impacting health.

Some variants, like those in cystic fibrosis and sickle cell disease, directly lead to illnesses. In conditions such as hypertension and diabetes, genomic variants may subtly influence overall risk. Healthcare experts are progressively integrating knowledge of patient's genomic variations into medical care, a practice referred to as genomic medicine.

Next-Generation Sequencing Technologies:

Human Genome Project: A Historic Milestone:

April 14, 2003 is one of the unforgettable and most important day for all Bioinformaticians, it is the day when the International Human Genome Sequencing Consortium announces the successful completion of the Human Genome Project. It is the blueprint that quite literally specifies how to build a human, even if the methods involved aren't completely comprehended yet. The Human Genome Project's primary achievement is sequencing the human genome for the first time, offered essential insights into the blueprint of humanity. This milestone has subsequently expedited the exploration of human biology and enhanced medical practices.

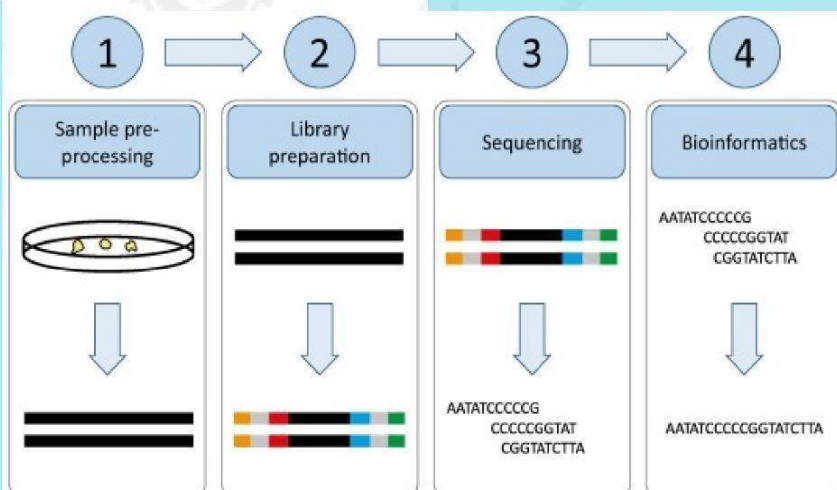


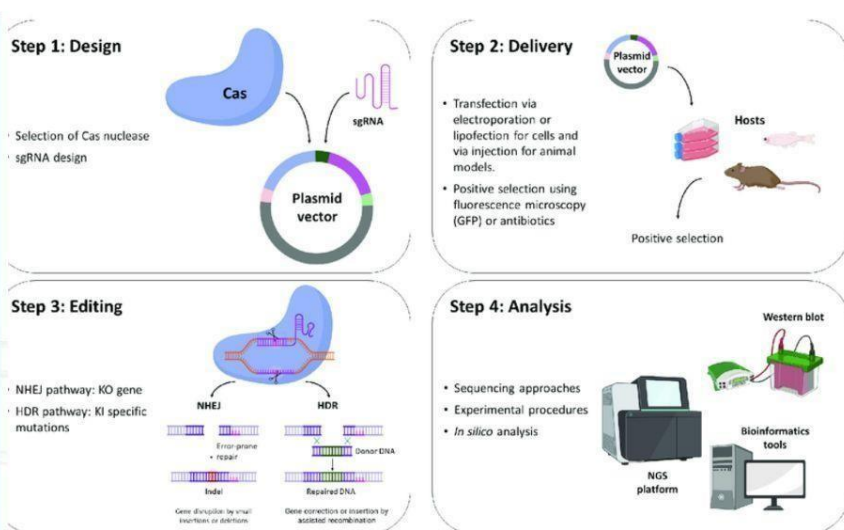
Fig. Next-Generation Sequencing Technologies workflow

VIVEKAA.A I YEAR M.SC BIOINFORMATICS & DATA SCIENCE	DECODING THE HUMAN BLUEPRINT: ADVANCES IN GENOMIC RESEARCH	CONTINUED
--	---	----------------

It's hard to believe that entire human genome can be sequenced within a single day, but it's possible with the help of Next Generation Sequencing (NGS) technology. In contrast to the previous Sanger sequencing technology, used to decipher the human genome, required over a decade to deliver the final draft. It is a massively parallel sequencing technology that offers ultra-high throughput, scalability, and speed. This technology is applied in sequencing of DNA and RNA and widely used in detection of variations and mutations which is applied in disease diagnosis, prognosis, therapeutic decision, and follow up of patients. The foundational process of next-generation sequencing entails the fragmentation of DNA/RNA, incorporation of adapters, sequencing the resultant libraries, and subsequent reassembly to generate a genomic sequence. This methodology, resembling capillary electrophoresis in principle, diverges significantly by orchestrating the simultaneous sequencing of myriad fragments in a massively parallel paradigm. This innovative approach not only accelerates the process but also elevates accuracy, all the while mitigating the overall cost of sequencing.

CRISPR-Cas9 and Genome Editing:

Genome editing stands as a sophisticated methodology that enables scientists to modify the DNA and RNA of diverse organisms ranging from prokaryotes and eukaryotes. This process results in transformative changes in physical attributes, such as eye colour, and modulates susceptibility to diseases. Researchers deploy an array of sophisticated technologies to execute this genome engineering. The CRISPR/Cas9 system is a research hotspot in genome editing and regulation.



CRISPR-Cas9 Technologies workflow

Clustered regularly interspaced palindromic repeats (CRISPR)/Cas9 is a gene-editing technology which plays a major role in development of biomedical research.

It facilitates the rectification of genomic errors and expeditious modulation of gene expression in cells and organisms, presenting a cost-effective and with notable simplicity. It's applications in the laboratory span from quickly creating cellular and animal models to conducting functional genomic screens and live imaging of the cellular genome. The applications of CRISPR-CAS9 is so wide.

It is adapted in activation and suppression of gene expression by attaching heterologous domains for altering epigenetic signatures on histone or transcriptional regulators. It is used in selective labelling of genome for monitoring dynamic chromatin processes.

This system has excellent therapeutic potential for curing various diseases with known genetic dysfunctions.

This leads to restoration of gene function or effectuate compensation for mutations within the genetic framework.

Future Prospects: Genomic Medicine

Since the conclusion of the Human Genome Project (HGP), there has been a substantial sense of optimism generated by genomic medicine. It assured earlier diagnosis and prevention of disease within the framework of precision medicine. Integrated omics uplifted personalised and precision medicine especially pharmacogenomics played a vital part in differentiating responders and non-responders to medicine in order to overcome adverse drug reactions and toxicity and achieve the optimum dose. Personalized medicine involves tailoring medical treatments based on individual patients' genetic background and disease status, selecting the most suitable treatment and therapeutic agent for each person. This method is widely used in diagnosis and treatment of cancer. The research is still in progress for infectious diseases, mendelian diseases and lot others. Genomic medicine is expected to attain its peak in forthcoming decade.

4. BOOK CHAPTER

INTERDISCIPLINARY PERSPECTIVES IN SYSTEMS BIOLOGY:FROM THEORY TO PRACTICE - ISBN Number: 978-93-340-3987-0

In my article "Metabolomics 101: Navigating the World of Small Molecules," I aimed to provide a comprehensive introduction to metabolomics and its significance in studying small molecules within biological systems. I began by defining metabolomics and emphasizing its role in understanding biochemical processes. The article explores the various analytical techniques employed in metabolomics research, including MS, NMR, and chromatography, which are crucial for identifying and quantifying metabolites with precision.

A significant portion of my article focused on the practical applications of metabolomics, particularly in disease diagnosis and personalized medicine. I discussed how metabolomics can reveal unique metabolic fingerprints associated with different health conditions, enabling early detection and targeted treatment strategies. Furthermore, I explored the intersection of metabolomics with drug discovery, highlighting its role in elucidating drug mechanisms and optimizing therapeutic interventions by targeting specific metabolic pathways. Metabolomics emerged as a powerful tool for enhancing pharmaceutical research and development.

I concluded the article by emphasizing the integrative nature of metabolomics within systems biology. I discussed how metabolomics synergizes with other omics disciplines, such as genomics, transcriptomics, and proteomics, to provide a holistic understanding of biological systems. By integrating multiple layers of biological information, metabolomics contributes to unraveling complex biological processes and networks. Overall, "Metabolomics 101: Navigating the World of Small Molecules" serves as a valuable resource for researchers and practitioners seeking to leverage metabolomics for advancing our understanding of biology and enhancing various applications in healthcare and drug development. Through my article, I aimed to contribute to the growing body of literature on metabolomics by providing a foundational overview accessible to both novice and seasoned researchers. I underscored the importance of metabolomics as a key player in systems biology, elucidating the intricate relationships between small molecules and biological phenomena. By presenting clear explanations and illustrative examples, my goal was to demystify metabolomics and inspire further exploration into the diverse applications of this transformative field.

4.1 Metabolomics 101: Navigating the World of Small Molecules

Navigating the World of Small Molecules:

Metabolomics, a growing field in the study of biological sciences, provides a broad view of the complex chemical processes in living organisms. In this article, we explore small molecules and their significant effects on health, disease, and drug discovery. Metabolomics helps scientists understand the detailed metabolic pathways and molecular patterns that define life. This article acts as a guide through the diverse aspects of metabolomics, from the tools used in analysis to its applications in diagnosing diseases, developing drugs, and understanding biological systems. Let us uncover the mysteries of metabolic patterns, examine analytical methods, discuss drug discovery, and explore how metabolomics fits into the broader study of biology. Come along as we unravel the role of small molecules in understanding the essence of life at a molecular level.

Tools of the Trade: Analytical Techniques in Metabolomics:

Metabolomics, a pivotal approach in unraveling disease phenotypes, delves into the intricate realm of small-molecule compounds such as lipids, amino acids, and organic acids. Employing a diverse array of analytical platforms including Gas Chromatography Mass Spectrometry (GC-MS), Nuclear Magnetic Resonance (NMR), Direct Infusion Mass Spectrometry (DI-MS), Two-dimensional Gas Chromatography Mass Spectrometry (GC × GC-MS), Capillary Electrophoresis Mass Spectrometry (CE-MS), and Liquid Chromatography Mass Spectrometry (LC-MS), researchers navigate the complex landscape of metabolites' diverse properties. Integration of these techniques augments coverage, recognizing that no single platform can comprehensively detect all metabolites. Integrated platforms facilitate sensitive and reliable detection of myriad metabolites in biofluid samples, propelling metabolomic research forward. Automation and machine learning, particularly in NMR and GC-MS, enhance efficiency. Notably, GC × GC-MS surpasses traditional methods in detecting metabolites, while GC-MS excels in identifying thermostable volatile compounds. Additionally, Gas Chromatography coupled to Time-of-Flight Mass Spectrometry (GC-ToF-MS) emerges prominently, especially in large-scale plant metabolomics, due to its efficient detection of a wide range of metabolites. The ongoing evolution of these platforms further integrates metabolomics into the realm of

systems biology, illuminating the intricate interplay of molecules within the body and offering invaluable insights into health and disease.

Metabolic Fingerprints: Using Metabolomics for Disease Diagnosis:

The rising demand for comprehensive understanding and detection of biological endpoints in human diseases has propelled the prominence of metabolomics, offering a promising avenue for disease diagnosis, prognosis, and therapeutic development. Metabolomics, along with related fields like metabonomics and metabolite profiling, entails the quantitative detection of multiple small molecule metabolites in biological systems. By combining metabolomics with genomics and proteomics, a deeper insight into biological systems at the molecular level, known as systems biology, is anticipated. This integration holds the potential for personalized medicine, tailoring treatments to individual metabolic profiles.

Metabolomics boasts the advantage of detecting numerous metabolites simultaneously, offering a sensitive measure of biological status. With approximately 3000 to 5000 detectable metabolites in the human body, changes in metabolite concentrations can signal perturbations in disease-related pathways. Factors such as diet, age, and lifestyle influence the metabolic profile, necessitating meticulous control or deconvolution to glean disease-specific information. Analytical techniques like nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS) are commonly employed in metabolomics studies due to their high reproducibility and sensitivity.

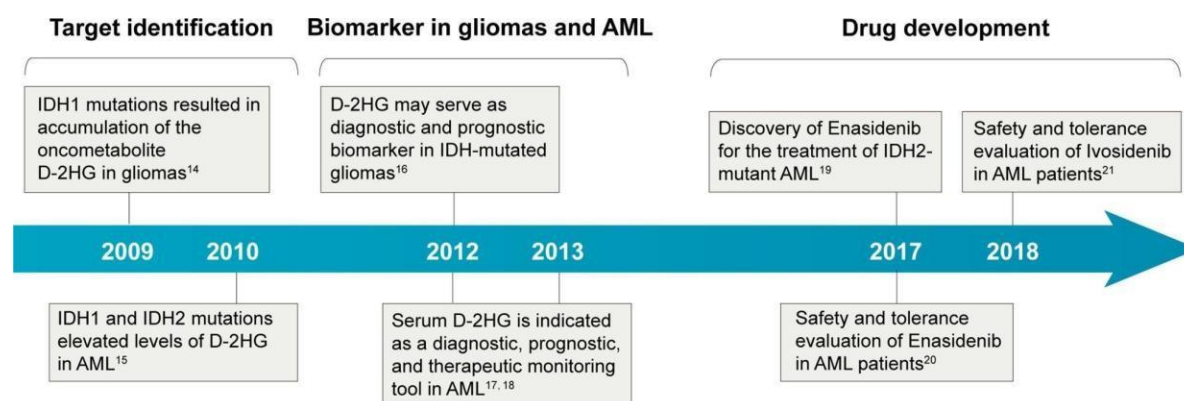
NMR spectroscopy, renowned for its ability to analyze multi-component mixtures with minimal sample preparation, provides detailed information on metabolite identity and quantity. Similarly, MS, with its intrinsic high sensitivity, is crucial for measuring metabolites in complex biofluids. Multivariate statistical analyses aid in data reduction, facilitating the differentiation of disease and control populations based on metabolite signals. These statistical methods also aid in identifying single biomarkers or sets of biomarkers for disease assessment, while validation ensures the reliability of findings, supported by a deeper biological understanding of disease processes. Metabolomics thus holds great promise in advancing diagnostic capabilities and enhancing our understanding of disease mechanisms.

In a recent case study, metabolomics played a pivotal role in elucidating the metabolic signatures of Alzheimer's disease (AD). Researchers utilised NMR spectroscopy and MS to analyse cerebrospinal fluid samples from AD patients and healthy controls. Multivariate statistical analyses revealed distinct metabolic profiles between the two groups, with alterations in key metabolites associated with neurotransmitter pathways, lipid metabolism, and oxidative stress. Furthermore, the identification of specific biomarkers, such as elevated levels of amyloid beta-related peptides and decreased levels of certain amino acids, provided valuable insights into the pathophysiology of AD and potential diagnostic markers. This study exemplifies the power of metabolomics in uncovering disease-specific metabolic alterations and highlights its potential for early detection and monitoring of neurodegenerative disorders.

Metabolomics in Drug Discovery: Targeting Metabolic Pathways:

Metabolomics, a burgeoning field in biology, offers a comprehensive look into metabolites and how they function in living systems. It's shedding light on the complex processes behind diseases like cancer, heart disease, and diabetes, providing crucial insights into how they work. By identifying disease biomarkers, metabolomics helps us understand diseases better and aids in developing new drugs.

In drug research, metabolomics is proving to be invaluable. It helps scientists grasp disease mechanisms, pinpoint drug targets, and understand how drugs work. Plus, it's speeding up the drug development process, making it more efficient and cost-effective. Take, for example, the development of drugs like Ivosidenib and Enasidenib for treating acute myeloid leukaemia (AML). Metabolomics identified a key metabolite, D-2HG, linked to the disease, leading to the development of these targeted therapies.



Drug research and development have become increasingly challenging and expensive over the years. With metabolomics, though, there's hope for making the process smoother. Metabolomics aids in screening potential drugs, assessing their safety, and monitoring adverse reactions, reducing the risk of costly failures. In recent years, it's also been crucial in optimizing the attributes of drugs and increasing their success rates. Overall, metabolomics is proving to be a game-changer in the world of drug discovery and development, offering a promising path forward in finding effective treatments for a range of diseases.

Integrative Approaches: Metabolomics in Systems Biology:

The emerging field of metabolomics offers novel techniques for acquiring comprehensive metabolic data, facilitating the investigation of microbial metabolomes and metabolic networks. Leveraging algorithms for metabolic network reconstruction and modelling, metabolomics contributes to the development of systems biology. Even qualitative models, constrained by stoichiometric principles, provide valuable insights into complex biochemical systems, serving as a foundation for quantitative models. Metabolomics data, akin to other 'omics' data, form a resource for hypothesis generation and testing, challenging conventional reductionist approaches. Metabolomics, intertwined with analytical technology, genomics, and computation, plays a pivotal role in systems biology. Advances in metabolomic data acquisition and analysis are crucial for understanding metabolic network organization and function. Integration of metabolomics into systems biology involves quantitative experimentation and mathematical modelling, facilitating holistic insights into complex biological systems. Moreover, metabolomics contributes to the discovery of biomarkers and disease status, with implications for biomedical research and clinical applications. Future developments in metabolomics will focus on enhancing instrument sensitivity, universality, and discrimination, alongside addressing challenges in metabolite identification and data deconvolution. Metabolomics stands poised at the forefront of systems biology, driving forward our understanding of biological systems and their intricate metabolic networks.

References:

1. Zhang A, Sun H, Wang P, Han Y, Wang X. Modern analytical techniques in metabolomics analysis. *Analyst*. 2012 Jan 21;137(2):293-300. doi: 10.1039/c1an15605e. Epub 2011 Nov 21. PMID: 22102985.
2. Gowda GA, Zhang S, Gu H, Asiago V, Shanaiah N, Raftery D. Metabolomics-based methods for early disease diagnostics. *Expert Rev Mol Diagn*. 2008 Sep;8(5):617-33. doi: 10.1586/14737159.8.5.617. PMID: 18785810; PMCID: PMC3890417.
3. Nath, D., & Chetia, D. (2024). Metabolomics: Special Emphasis on Basic Drug Discovery and Development. *IntechOpen*. doi: 10.5772/intechopen.112969
4. Douglas B Kell, Metabolomics and systems biology: making sense of the soup, *Current Opinion in Microbiology*, Volume 7, Issue 3, 2004, Pages 296-307, ISSN 1369-5274, <https://doi.org/10.1016/j.mib.2004.04.012>.
5. Huanhuan Pang, Zeping Hu, Metabolomics in drug research and development: The recent advances in technologies and applications, *Acta Pharmaceutica Sinica B*, Volume 13, Issue 8, 2023, Pages 3238-3251, ISSN 2211-3835, <https://doi.org/10.1016/j.apsb.2023.05.021>.

Vivekaa A

by OfficialTurnitin

Submission date: 29-Apr-2024 12:45AM (UTC+0900)

Submission ID: 2362321588

File name: Final-article-21.docx (289.83K)

Word count: 1116

Character count: 7407

Navigating the World of Small Molecules:

Metabolomics, a growing field in the study of biological sciences, provides a broad view of the complex chemical processes in living organisms. In this article, we explore small molecules and their significant effects on health, disease, and drug discovery. Metabolomics helps scientists understand the detailed metabolic pathways and molecular patterns that define life. This article acts as a guide through the diverse aspects of metabolomics, from the tools used in analysis to its applications in diagnosing diseases, developing drugs, and understanding biological systems. Let us uncover the mysteries of metabolic patterns, examine analytical methods, discuss drug discovery, and explore how metabolomics fits into the broader study of biology. Come along as we unravel the role of small molecules in understanding the essence of life at a molecular level.

Tools of the Trade: Analytical Techniques in Metabolomics:

Metabolomics, a pivotal approach in unraveling disease phenotypes, delves into the intricate realm of small-molecule compounds such as lipids, amino acids, and organic acids. Employing a diverse array of analytical platforms including (GC-MS), (NMR), (DI-MS), (GC $\times$ GC-MS), (CE-MS), and (LC-MS), researchers navigate the complex landscape of metabolites' diverse properties. Integration of these techniques augments coverage, recognizing that no single platform can comprehensively detect all metabolites. Integrated platforms facilitate sensitive and reliable detection of myriad metabolites in biofluid samples, propelling metabolomic research forward. Automation and machine learning, particularly in NMR and GC-MS, enhance efficiency. Notably, GC $\times$ GC-MS surpasses traditional methods in detecting metabolites, while GC-MS excels in identifying thermostable volatile compounds. Additionally, (GC-ToF-MS) emerges prominently, especially in large-scale plant metabolomics, due to its efficient detection of a wide range of metabolites. The ongoing evolution of these platforms further integrates metabolomics into the realm of systems biology, illuminating the intricate interplay of molecules within the body and offering invaluable insights into health and disease.

Metabolic Fingerprints: Using Metabolomics for Disease Diagnosis:

The rising demand for comprehensive understanding and detection of biological endpoints in human diseases has propelled the prominence of metabolomics, offering a promising avenue for disease diagnosis, prognosis, and therapeutic development. Metabolomics, along with related fields like metabonomics and metabolite profiling, entails the quantitative detection of multiple small molecule metabolites in biological systems. By combining metabolomics with genomics and proteomics, a deeper insight into biological systems at the molecular level, known as systems biology, is anticipated. This integration holds the potential for personalized medicine, tailoring treatments to individual metabolic profiles.

Metabolomics boasts the advantage of detecting numerous metabolites simultaneously, offering a sensitive measure of biological status. With approximately 3000 to 5000 detectable metabolites in the human body, changes in metabolite concentrations can signal perturbations in disease-related pathways. Factors such as diet, age, and lifestyle influence the metabolic profile, necessitating meticulous control or deconvolution to glean disease-specific information. Analytical techniques like (NMR) (MS) are commonly employed in metabolomics studies due to their high reproducibility and sensitivity.

NMR spectroscopy, renowned for its ability to analyze multi-component mixtures with minimal sample preparation, provides detailed information on metabolite identity and quantity. Similarly, MS, with its intrinsic high sensitivity, is crucial for measuring metabolites in complex biofluids. Multivariate statistical analyses aid in data reduction, facilitating the differentiation of disease and control populations based on metabolite signals. These statistical methods also aid in identifying single biomarkers or sets of biomarkers for disease assessment, while validation ensures the reliability of findings, supported by a deeper biological understanding of disease processes. Metabolomics thus holds great promise in advancing diagnostic capabilities and enhancing our understanding of disease mechanisms.

In a recent case study, metabolomics played a pivotal role in elucidating the metabolic signatures of (AD). Researchers utilised NMR spectroscopy and MS to analyse cerebrospinal fluid samples from AD patients and healthy controls. Multivariate statistical analyses revealed distinct metabolic profiles between the two groups, with alterations in key metabolites associated with neurotransmitter pathways, lipid metabolism, and oxidative stress. Furthermore, the identification of specific biomarkers, such as elevated levels of amyloid beta-related peptides and decreased levels of certain amino acids, provided valuable insights into the pathophysiology of AD and potential diagnostic markers. This study exemplifies the power of metabolomics in uncovering disease-specific metabolic alterations and highlights its potential for early detection and monitoring of neurodegenerative disorders.

Metabolomics in Drug Discovery: Targeting Metabolic Pathways:

Metabolomics, a burgeoning field in biology, offers a comprehensive look into metabolites and how they function in living systems. It's shedding light on the complex processes behind diseases like cancer, heart disease, and diabetes, providing crucial insights into how they work. By identifying disease biomarkers, metabolomics helps us understand diseases better and aids in developing new drugs.

In drug research, metabolomics is proving to be invaluable. It helps scientists grasp disease mechanisms, pinpoint drug targets, and understand how drugs work. Plus, it's speeding up the drug development process, making it more efficient and cost-effective. Take, for example, the development of drugs like Ivosidenib and Enasidenib for treating (AML). Metabolomics identified a key metabolite, D-2HG, linked to the disease, leading to the development of these targeted therapies.

Drug research and development have become increasingly challenging and expensive over the years. With metabolomics, though, there's hope for making the process smoother. Metabolomics aids in screening potential drugs, assessing their safety, and monitoring adverse reactions, reducing the risk of costly failures. In recent years, it's also been crucial in optimizing the attributes of drugs and increasing their success rates. Overall, metabolomics is proving to be a game-changer in the world of drug discovery and development, offering a promising path forward in finding effective treatments for a range of diseases.

Integrative Approaches: Metabolomics in Systems Biology:

The emerging field of metabolomics offers novel techniques for acquiring comprehensive metabolic data, facilitating the investigation of microbial metabolomes and metabolic networks. Leveraging algorithms for metabolic network reconstruction and modelling, metabolomics contributes to the development of systems biology. Even qualitative models, constrained by stoichiometric principles, provide valuable insights into complex biochemical systems, serving as a foundation for quantitative models. Metabolomics data, akin to other 'omics' data, form a resource for hypothesis generation and testing, challenging conventional reductionist approaches. Metabolomics, intertwined with analytical technology, genomics, and computation, plays a pivotal role in systems biology. Advances in metabolomic data acquisition and analysis are crucial for understanding metabolic network organization and function. Integration of metabolomics into systems biology involves quantitative experimentation and mathematical modelling, facilitating holistic insights into complex biological systems. Moreover, metabolomics contributes to the discovery of biomarkers and disease status, with implications for biomedical research and clinical applications. Future developments in metabolomics will focus on enhancing instrument sensitivity, universality, and discrimination, alongside addressing challenges in metabolite identification and data deconvolution. Metabolomics stands poised at the forefront of systems biology, driving forward our understanding of biological systems and their intricate metabolic networks.

ORIGINALITY REPORT

5%

SIMILARITY INDEX

2%

INTERNET SOURCES

5%

PUBLICATIONS

1%

STUDENT PAPERS

PRIMARY SOURCES

1

GA Nagana Gowda, Shucha Zhang, Haiwei Gu, Vincent Asiago, Narasimhamurthy Shanaiah, Daniel Raftery. "Metabolomics-based methods for early disease diagnostics", Expert Review of Molecular Diagnostics, 2014
Publication

2%

2

www.tandfonline.com
Internet Source

2%

Exclude quotes Off
Exclude bibliography On

Exclude matches Off

5. NATIONAL CONFERENCE ON TRANSFORMING HEALTHCARE: ADVANCES IN BIOINFORMATICS AND BIOLOGICAL SCIENCES (ABBS - 2024)

As an organizing member of the "National Conference on Transforming Healthcare: Advances in Bioinformatics and Biological Sciences (ABBS-2024)," I am delighted to provide a detailed overview of this influential event. Hosted by the Department of Bioinformatics and Data Science at Sathyabama Institute of Science and Technology, ABBS-2024 gathered scientists, researchers, and students to explore the latest innovations in bioinformatics and computational biology.

ABBS-2024 was carefully structured to delve into critical areas driving healthcare transformation, including genomics, structural bioinformatics, biotechnology applications, and the integration of artificial intelligence in healthcare. Each session and workshop was thoughtfully curated to foster intellectual discourse and encourage interdisciplinary collaborations. The conference highlighted the pivotal role of bioinformatics in revolutionizing healthcare, from personalized medicine approaches to innovative drug discovery strategies.

Throughout the event, attendees participated in engaging presentations, interactive discussions, and poster sessions facilitated by notable keynote speakers and experts from academia and industry. The exchange of ideas and experiences among participants created a dynamic atmosphere of learning and innovation. The conference emphasized practical applications and real-world impact, showcasing how bioinformatics tools and computational techniques are reshaping disease diagnosis, treatment, and prevention.

ABBS-2024 also addressed ethical considerations and societal implications, ensuring that scientific advancements are guided by responsible practices and informed decision-making. The event celebrated scientific excellence while emphasizing the importance of inclusive research practices and knowledge dissemination for the benefit of society. It provided a valuable platform for young researchers and students to showcase their work, receive feedback, and establish connections for future collaborations. In conclusion, ABBS-2024 was a resounding success, driven by the collective dedication of participants towards advancing healthcare through bioinformatics and biological sciences.



SATHYABAMA

Institute of Science and Technology (Deemed to be University)
Accredited A++ by NAAC | 12B Status by UGC | Approved by AICTE



Certificate of Appreciation

This is to certify that

VIVEKAA .A

has demonstrated outstanding commitment and leadership as an **Student Organizer** for the successful execution of the event "**National Conference on Transforming Healthcare: Advances in Bioinformatics & Biological Sciences (ABBS 2024)**" by the **Department of Bioinformatics** from **13/02/2024 to 15/02/2024** at **Sathyabama Institute of Science & Technology**.

Dr. G.SUNDARI
Dean
Administration

Dr. REKHA CHAKRAVARTHI
Dean
Science & Humanities

Dr. H JEMMY CHRISTY
Head
Department of Bioinformatics

Dr. SWETHA SUNKAR
Organizing
secretary

6.MATLAB SELF PACED ONLINE COURSES

Over the course of my professional development, I have undertaken and successfully completed a rigorous series of MATLAB courses designed to enhance my proficiency in programming and data analysis. This report provides a comprehensive overview of the key concepts and skills acquired from each course, showcasing my expertise in MATLAB programming and its diverse applications.

6.1 COURSES COMPLETED:

Image Processing with MATLAB

The Image Processing with MATLAB course provided a comprehensive foundation in image processing techniques essential for tasks such as medical image analysis and computer vision applications. Throughout this course, I gained expertise in image enhancement, filtering, segmentation, and feature extraction using MATLAB's powerful toolset. These skills are crucial for analyzing and manipulating digital images effectively, enabling applications like medical diagnosis, object recognition, and satellite image processing. By completing this course, I developed the capability to apply advanced image processing algorithms, optimize image quality, and extract meaningful information from visual data using MATLAB, contributing to a deeper understanding of image-based tasks in various domains.

Signal Processing with MATLAB

Mastering Signal Processing with MATLAB equipped me with advanced knowledge in analyzing and manipulating signals crucial for various engineering disciplines. This course delved into Fourier transforms, signal filtering techniques, and spectral analysis using MATLAB's signal processing capabilities. Through practical exercises and projects, I honed my skills in processing sensor data, audio signals, and communication signals. This course enabled me to apply mathematical transformations to signals, extract valuable information, and make data-driven decisions. With proficiency in signal processing using MATLAB, I gained the ability to solve complex engineering problems, optimize signal quality, and develop

innovative solutions for real-world applications such as telecommunications, radar systems, and biomedical signal analysis.

Reinforcement Learning

The Reinforcement Learning course provided hands-on experience in implementing reinforcement learning algorithms using MATLAB, facilitating the development of intelligent decision-making systems and autonomous agents. Through this course, I learned foundational concepts like Q-learning and policy gradient methods, which are essential for training agents to interact with environments and learn optimal behaviors. By applying reinforcement learning techniques in MATLAB, I gained practical skills in solving sequential decision-making problems and optimizing agent performance in dynamic environments. This course prepared me to develop autonomous systems capable of learning from experience and adapting to changing conditions, with applications in robotics, game playing, and optimization tasks.

MATLAB for Data Processing and Visualization

MATLAB for Data Processing and Visualization offered advanced techniques for efficiently handling and analyzing complex datasets using MATLAB. This course focused on data preprocessing, statistical analysis, and creating insightful visualizations to extract meaningful insights from data. By leveraging MATLAB's capabilities, I gained proficiency in processing large datasets, performing data cleaning and transformation, and applying statistical methods for hypothesis testing and regression analysis. The skills acquired in this course are crucial for interpreting and presenting data effectively, enabling informed decision-making in scientific research, engineering, and business analytics.

Solving Nonlinear Equations with MATLAB

In the Solving Nonlinear Equations with MATLAB course, I acquired expertise in numerically solving complex nonlinear equations commonly encountered in scientific computing and engineering simulations. This course covered a range of numerical methods including Newton-Raphson and optimization-based approaches implemented using MATLAB's computational tools. By mastering

nonlinear equation solving techniques, I developed the ability to model and simulate nonlinear systems accurately, facilitating the analysis and optimization of engineering designs and mathematical models. This course enhanced my problem-solving skills and computational capabilities, enabling me to tackle challenging nonlinear problems in various domains effectively.

Optimization Onramp

The Optimization Onramp course introduced fundamental optimization methods and algorithms within MATLAB, focusing on linear programming and constrained optimization techniques. Through this course, I gained practical skills in formulating and solving optimization problems, optimizing resource allocation, and maximizing objective functions subject to constraints. By leveraging MATLAB's optimization toolbox, I learned to apply mathematical optimization techniques to real-world problems in engineering, operations research, and machine learning. This course provided a strong foundation in optimization principles, empowering me to develop efficient algorithms and solve complex optimization problems efficiently using MATLAB.

MATLAB Programming Techniques

The MATLAB Programming Techniques course expanded my proficiency in MATLAB programming, emphasizing advanced programming techniques such as code optimization, algorithm design, and software development practices. This course equipped me with the skills to write modular, efficient, and maintainable code suitable for complex applications. By applying programming best practices in MATLAB, I enhanced my ability to design scalable algorithms, implement data structures, and optimize performance. The knowledge gained from this course is invaluable for developing robust and scalable MATLAB applications across diverse domains, improving productivity and code quality in software development projects.

Image Processing Onramp

The Image Processing Onramp course served as an introductory exploration of image processing concepts and MATLAB's image processing toolbox. This course provided foundational knowledge in image representation, basic operations, and

introductory filtering techniques using MATLAB. By completing this course, I gained a solid understanding of image processing fundamentals, setting the stage for more advanced exploration of image analysis and computer vision tasks in MATLAB. The skills acquired in this course are essential for anyone working with digital images, from basic image manipulation to advanced feature extraction and pattern recognition using MATLAB's comprehensive image processing capabilities.

MATLAB Fundamentals

The MATLAB Fundamentals course provided a robust foundation in MATLAB programming, covering essential syntax, data types, and programming constructs. Through this course, I gained proficiency in variables, arrays, control flow (loops, conditionals), and function writing in MATLAB. This foundational knowledge enabled me to develop and execute MATLAB scripts, perform basic computations, and organize code effectively. By mastering MATLAB fundamentals, I acquired the skills necessary to implement algorithms, solve mathematical problems, and develop simple applications using MATLAB's powerful programming environment.

Statistics Onramp

The Statistics Onramp course introduced statistical analysis techniques using MATLAB, focusing on descriptive statistics, hypothesis testing, regression analysis, and data visualization. By completing this course, I gained practical skills in analyzing data, making data-driven decisions, and communicating insights through statistical methods and graphical representations. This course equipped me with the ability to perform hypothesis tests, fit regression models, and interpret statistical results using MATLAB. The knowledge gained in statistical analysis is invaluable for drawing meaningful conclusions from data and supporting decision-making processes in research, engineering, and business analytics.

Machine Learning with MATLAB

The Machine Learning with MATLAB course provided comprehensive training in applying machine learning algorithms and techniques using MATLAB. Throughout this course, I gained insights into supervised and unsupervised learning methods, model training, validation, and deployment of machine learning models. By leveraging MATLAB's machine learning toolbox, I acquired practical skills in

developing predictive models, analyzing data patterns, and evaluating model performance. This course prepared me to leverage machine learning algorithms for classification, regression, clustering, and other tasks, enabling data-driven decision-making and pattern recognition applications.

Deep Learning with MATLAB

Exploring Deep Learning with MATLAB expanded my knowledge of deep neural networks, convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transfer learning techniques. This course equipped me with the skills to design, train, and deploy deep learning models using MATLAB's deep learning toolbox. By completing practical exercises and projects, I gained expertise in image classification, object detection, and natural language processing tasks. Deep Learning with MATLAB prepared me to leverage cutting-edge deep learning techniques for solving complex problems in computer vision, speech recognition, and sequence modeling applications.

Machine Learning Onramp

The Machine Learning Onramp course provided a foundational understanding of machine learning principles and applications using MATLAB. This introductory course covered key concepts such as feature selection, model training, and evaluation of machine learning algorithms. By completing this course, I gained insight into the workflow of machine learning projects, from data preprocessing to model deployment. This foundational knowledge in machine learning enabled me to explore more advanced techniques and algorithms for solving classification, regression, and clustering problems using MATLAB's machine learning tools.

Deep Learning Onramp

Similar to the Machine Learning Onramp, the Deep Learning Onramp focused specifically on deep learning methodologies using MATLAB. This introductory course covered advanced deep learning concepts, including neural network architectures, training strategies, and transfer learning. By completing this course, I gained practical skills in designing and implementing deep learning models for image recognition, natural language processing, and time series analysis tasks.

Deep Learning Onramp provided the foundation necessary to explore advanced topics in deep learning and apply cutting-edge techniques using MATLAB.

Computer Vision Onramp

The Computer Vision Onramp course equipped me with practical knowledge of computer vision techniques and algorithms using MATLAB. This introductory course covered image feature extraction, object detection, and image classification using MATLAB's computer vision toolbox. By completing this course, I gained the skills necessary to develop applications for image-based tasks, such as automated inspection, facial recognition, and augmented reality. Computer Vision Onramp prepared me to leverage MATLAB's capabilities for solving real-world computer vision problems and developing innovative image processing applications.

Object-Oriented Programming Onramp

The Object-Oriented Programming (OOP) Onramp introduced fundamental principles of object-oriented programming in the context of MATLAB. This course covered classes, inheritance, polymorphism, and encapsulation, emphasizing code modularity and reusability. By completing this course, I gained proficiency in designing modular, maintainable, and extensible code using OOP concepts in MATLAB. This foundational knowledge in OOP enhanced my ability to organize complex projects, improve code structure, and collaborate effectively in software development teams.

MATLAB Onramp

The MATLAB Onramp served as a beginner-level introduction to MATLAB programming, covering basic syntax, data types, and simple programming tasks. This course provided essential skills for navigating the MATLAB environment, writing scripts, and performing basic computations. By completing this course, I acquired the foundational knowledge necessary to start exploring more advanced MATLAB functionalities and applications. MATLAB Onramp served as a stepping stone for further learning and development in MATLAB programming and computational techniques.

Solving ODE with MATLAB

The Solving Ordinary Differential Equations (ODEs) with MATLAB course focused on numerical methods for solving ODEs using MATLAB's computational tools. This course covered techniques such as Euler's method, Runge-Kutta methods, and boundary value problem solvers. By completing this course, I acquired the skills to model dynamic systems, simulate physical processes, and solve engineering problems involving differential equations. The ability to solve ODEs numerically is essential for mathematical modeling, system analysis, and simulation in various scientific and engineering fields.

Introduction to Linear Algebra with MATLAB

The Introduction to Linear Algebra with MATLAB course provided a comprehensive understanding of linear algebra concepts and their implementation in MATLAB. Topics included matrix operations, eigenvalues, eigenvectors, and matrix decompositions (LU, QR). By completing this course, I gained proficiency in using MATLAB for linear algebra computations, which are foundational for scientific computing, data analysis, and machine learning algorithms. This course equipped me with essential mathematical skills for solving systems of linear equations, analyzing geometric transformations, and performing matrix computations efficiently using MATLAB.

Signal Processing Onramp

The Signal Processing Onramp course introduced fundamental concepts and techniques of signal processing using MATLAB. Topics covered included sampling, filtering, Fourier analysis, and spectral analysis. By completing this course, I gained a solid understanding of signal processing fundamentals, preparing me for more advanced signal analysis tasks. This course provided the knowledge and skills necessary to process and analyze various types of signals, such as audio signals, biomedical signals, and communication signals, using MATLAB's signal processing toolbox.

6.2 CONCLUSION:

The completion of these MATLAB courses has significantly enhanced my programming capabilities and analytical skills, enabling me to tackle complex problems in engineering, data science, and research with confidence. I am well-

prepared to leverage MATLAB's powerful toolset to develop innovative solutions and contribute effectively to projects requiring advanced computational and analytical expertise.



Course Completion Certificate

Vivekaa A

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

MATLAB for Data Processing and Visualization


DIRECTOR, TRAINING SERVICES

5 March 2024



Course Completion Certificate

Vivekaa A

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

Optimization Onramp


DIRECTOR, TRAINING SERVICES

27 March 2024



Course Completion Certificate

Vivekaa A

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

Image Processing Onramp


DIRECTOR, TRAINING SERVICES

5 March 2024



Course Completion Certificate

Vivekaa A

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

Object-Oriented Programming Onramp


DIRECTOR, TRAINING SERVICES

19 December 2023



Course Completion Certificate

Vivekaa A

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

Signal Processing Onramp


DIRECTOR, TRAINING SERVICES

2 April 2024



Course Completion Certificate

Vivekaa A

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

Image Processing with MATLAB


DIRECTOR, TRAINING SERVICES

22 March 2024



Course Completion Certificate

Vivekaa A

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

Signal Processing with MATLAB


DIRECTOR, TRAINING SERVICES

2 April 2024



Course Completion Certificate

Vivekaa A

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

Reinforcement Learning Onramp


DIRECTOR, TRAINING SERVICES

21 February 2024



Course Completion Certificate

Vivekaa A

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

Solving Nonlinear Equations with MATLAB

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

DIRECTOR, TRAINING SERVICES

6 March 2024



Course Completion Certificate

Vivekaa A

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

MATLAB Fundamentals

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

DIRECTOR, TRAINING SERVICES

21 March 2024



Course Completion Certificate

Vivekaa A

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

MATLAB Programming Techniques


DIRECTOR, TRAINING SERVICES

5 March 2024



Course Completion Certificate

Vivekaa A

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

Statistics Onramp


DIRECTOR, TRAINING SERVICES

4 March 2024



Course Completion Certificate

Vivekaa A

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

Machine Learning with MATLAB


DIRECTOR, TRAINING SERVICES

19 January 2024



Course Completion Certificate

Vivekaa A

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

Deep Learning with MATLAB


DIRECTOR, TRAINING SERVICES

16 February 2024



Course Completion Certificate

Vivekaa A

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

Machine Learning Onramp



DIRECTOR, TRAINING SERVICES

19 December 2023



Course Completion Certificate

Vivekaa A

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

Computer Vision Onramp



DIRECTOR, TRAINING SERVICES

29 December 2023



Course Completion Certificate

Vivekaa A

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

MATLAB Onramp



DIRECTOR, TRAINING SERVICES

16 November 2023



Course Completion Certificate

Vivekaa A

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

Deep Learning Onramp



DIRECTOR, TRAINING SERVICES

17 January 2024



Course Completion Certificate

Vivekaa A

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

Solving Ordinary Differential Equations with MATLAB



DIRECTOR, TRAINING SERVICES

2 April 2024



Course Completion Certificate

Vivekaa A

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

Introduction to Linear Algebra with MATLAB



DIRECTOR, TRAINING SERVICES

11 April 2024