

**MINI PROJECT REPORT
(2021-2022)**

On

**“Erudition in Artificial Intelligence
&Machine Learning”**

Submitted by

**Isha Singh
(191500355)**

**Abhishek
Chaubey
(191500028)**

**Department of Computer Engineering & Applications
Institute of Engineering & Technology**



**GLA University
Mathura- 281406, INDIA**



**Department of computer Engineering and Applications
GLA University, Mathura**

**17 km. Stone NH#2, Mathura-Delhi Road, P.O. – Chaumuha,
Mathura – 281406**

Declaration

We hereby declare that the work which is being presented in the Mini-Project **“Erudition in Artificial Intelligence & Machine Learning”**, in partial fulfillment of the requirements for Mini-project viva-voce, is an authentic record of our own work carried by the team members under the supervision of our mentor Ms. Ruchi Gupta.

Group Members:

- Abhishek Kumar Chaubey (191500028)
- Isha Singh (191500355)

Course: B. Tech (Computer Science and Engineering)

Year: 3rd

Semester: 5th

Supervised By:
Ms. Ruchi Gupta, Technical Trainer

GLA University, Department of Computer Engineering & Application



**Department of computer Engineering and Applications
GLA University, Mathura**

**17 km. Stone NH#2, Mathura-Delhi Road, P.O. – Chaumuha,
Mathura – 281406**

Certificate

This is to certify that the above statements made by the candidates are correct to the best of my/our knowledge and belief.

Supervisor

Ms. Ruchi Gupta

Technical Trainer

Dept of CEA, GLA University

Project Coordinator

(Ms. Ruchi Gupta)

Program Coordinator

(Mr. Shashi Shekar)

About the Project

As technology advances internet is becoming ever popular. Website has become an essential part, and if we see the use of online learning websites the growth of such websites is observed at its peak. Online education is becoming more and more popular amongst students around the world and it's very easy for one to clear their doubts in one go. Our project "***Erudition in Artificial Intelligence & Machine Learning***" is about making such website blog in which the study material related to the Data Science i.e., Artificial Intelligence and machine learning, Statistics, Data Visualization will be provided at one platform. On this project we will develop a public website accessible to all in which the content will be provided in most simplified manner and that will be easy understandable by users.

Motivation

This website will be very useful in future for all users and we can grow this website as the trend of Data Science will never be down and the upcoming generation will also be curious to know about the new things and we will always try to update this website in order to give the best content to our user.

Requirements

a). Software Requirements:

- Technology Implemented: Full Stack Web Development
- Languages/Technologies Used: HTML, CSS, JavaScript, jQuery
- IDE Used: Visual Studio Code
- Web Browser: Google Chrome
- Website content is about core Data Science (*i.e., Artificial Intelligence and machine learning, Statistics, Data Visualization*)

α. **GitHub:** GitHub is a code hosting platform for version control and collaboration. It lets you and others work together on projects from anywhere. GitHub Repository: A GitHub repository can be used to store a development project. It can contain folders and any type of files (*HTML, CSS, JavaScript, Documents, Data, Images*). A GitHub repository should also include a license file and a README file about the project. A GitHub repository can also be used to store ideas, or any resources that you want to share.

β. **Visual Studio Code:** Visual Studio Code is a free source-code editor made by Microsoft for Windows, Linux and macOS. [7] Features include support for debugging, syntax highlighting, intelligent code completion, snippets, code refactoring, and embedded Git. Users can change the theme, keyboard shortcuts, preferences, and install extensions that add additional functionality. Microsoft has released Visual Studio Code's source code on the VS Code repository of GitHub.com, under the permissive MIT License, while the compiled binaries are freeware.

b). Hardware Requirements:

- Processor Required: Intel i5
- Operating System: Windows 10
- RAM: 8GB
- Hardware Devices: Computer System
- Hard Disk: 256GB

Acknowledgement

We thank the almighty for giving us the courage and perseverance in completing the project. This project itself is an acknowledgement for all those people who have given us their heartfelt co-operation in making this project a grand success. We extend our sincere thanks to **Ms. Ruchi Gupta**, Technical Trainer at “GLA University, Mathura” for providing his valuable guidance at every stage of this project work. We are profoundly grateful towards the unmatched services rendered by her. And last but not least, we would like to express our deep sense of gratitude and earnest thanks giving to our dear parents for their moral support and heartfelt cooperation in doing the main project.

Erudition in Artificial Intelligence & Machine Learning

Abstract

As the name suggests, our project is all about an educational website which facilitates you to reduce the work load and make work simple and efficient. Our website is one way of giving users fast service because instead of going through every website to get understandable content is difficult, so users can surf through our site here they will have a quite good experience as all content is provided at one place topic wise. The limited time and resources have restricted us to incorporate, but in this project utmost care will be taken to make the system efficient, user friendly and attractive.

Contents

Acknowledgment08

Abstract09

1. Introduction:

 Introduction to Full stack12-13

 Pre-requisites.....14

2. Technologies Used:

 HTML.....15

 CSS16

3. List of Figures.20-26

4. Software Testing..........27-30

5. Conclusion 31

6.Bibliography.......... 32

Introduction

Front-end development refers to creating and maintaining the part of the website that users directly interact with. Front-end developers need to create a website's entire interface and user experience, as well as its design and overall look. They usually work with HTML, JavaScript and CSS languages to create a basic layout of the website and then add various visual elements to improve its aesthetic quality. Some of the most commonly performed tasks for a front-end developer include:

- Optimizing the user's experience on a website and making sure it's not interrupted by any design or functionality issues
- Creating rough concepts and using HTML, CSS and JavaScript to materialize them
- Developing an easy-to-use and intuitive user interface and gradually improving it based on user feedback
- Adapting a website's design to look and function properly on mobile devices
- Making sure that a website looks and functions according to the same parameters regardless of the internet browser that the end-user chooses to open it with
- Helping with organizing and managing the entire software workflow
- Constantly testing the website's front end for ease of use and potential errors and optimizing it to ensure a smoother user experience
- Considering SEO practices when creating a website's front end to ensure that it's correctly indexed by search engines and users have easier access to it.

HTML

Hypertext Markup Language (HTML) is a computer language that makes up most web pages and online applications. A hypertext is a text that is used to reference other pieces of text, while a markup language is a series of markings that tells web servers the style and structure of a document. HTML is not considered a programming language as it can't create dynamic functionality. Instead, with HTML, web users can create and structure sections, paragraphs, and links using elements, tags, and attributes.

CSS

Cascading Style Sheets, fondly referred to as CSS, is a simple design language intended to simplify the process of making web pages presentable.

CSS handles the look and feel part of a web page. Using CSS, you can control the color of the text, the style of fonts, the spacing between paragraphs, how columns are sized and laid out, what background images or colors are used, layout designs, variations in display for different devices and screen sizes as well as a variety of other effects.

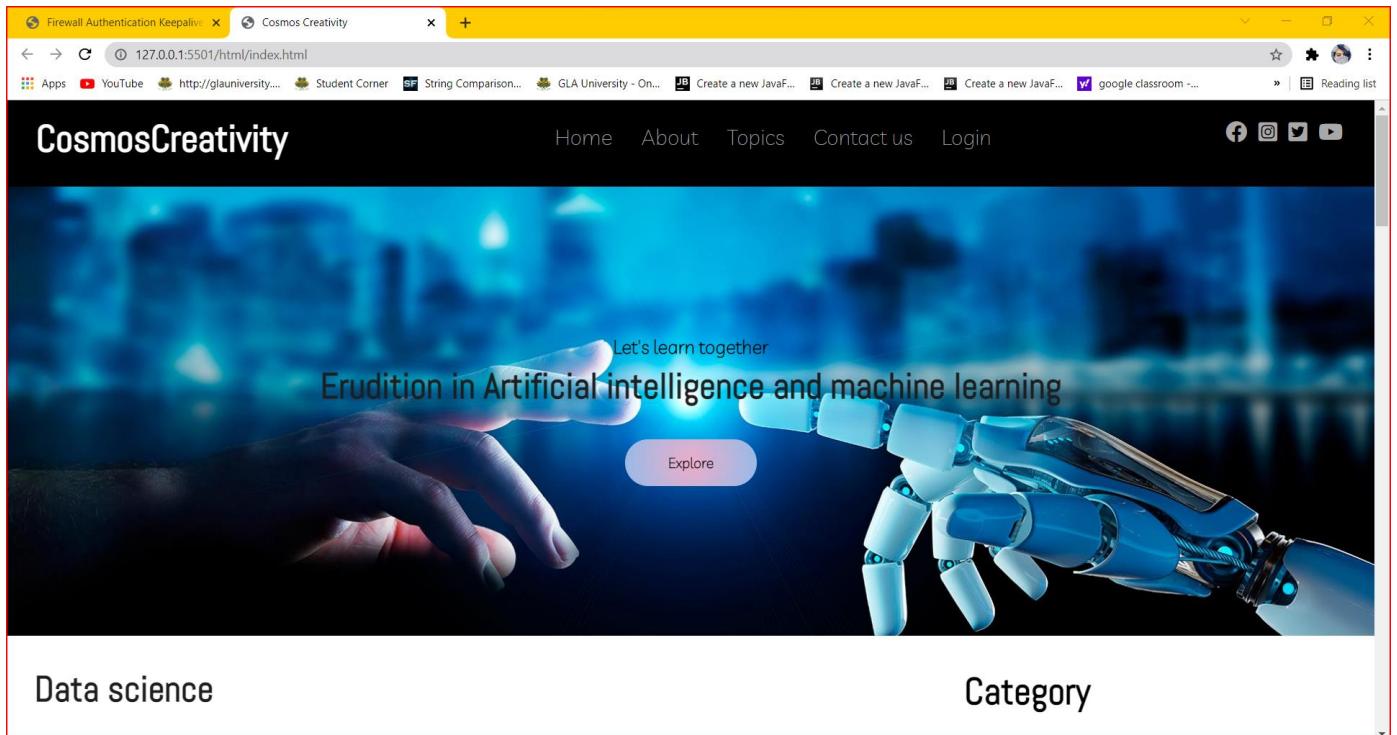
CSS is easy to learn and understand but it provides powerful control over the presentation of an HTML document. Most commonly, CSS is combined with the markup languages HTML or XHTML.

Pre-requisite

Hands-on knowledge of JavaScript, HTML and CSS is essential before working on the concepts for making of webpages. Make sure that you have the browser or chrome installed and running before opening website.

SNAPSHOTS FROM THE PROJECT

1. Home Page



What is Data Science?

Data science is subject field that combines programming skills, and knowledge of mathematics, statistics and domain expertise to extract meaningful insights from data. It encompasses preparing data for analysis, including cleansing, aggregating, and manipulating the data to perform advanced data analysis. Analytic applications and data scientists can then review the results to uncover patterns and enable business leaders to draw informed insights. Data

Applications of Data science

5 Application Areas of Data Science

Their are uncountable applications of Data Science but here are some main applications

1. Internet Search
2. Targeted Advertising
3. Website Recommendations
4. Advanced Image Recognition
5. Speech Recognition
6. Banking
7. e-Commerce
8. Finance

Popular Tags

- Regression
- Supervised ML
- Clustering
- Reinforcement-learning
- Classification
- Hypothesis
- Accuracy

About us

We provide the whole scenario of data science at one platform. Data Science comprises of different branch consist of Artificial Intelligence, Machine Learning, Statistics, Data Processing, Data Visualization, Software Engineering, So,

Reviews

Your reviews

write here

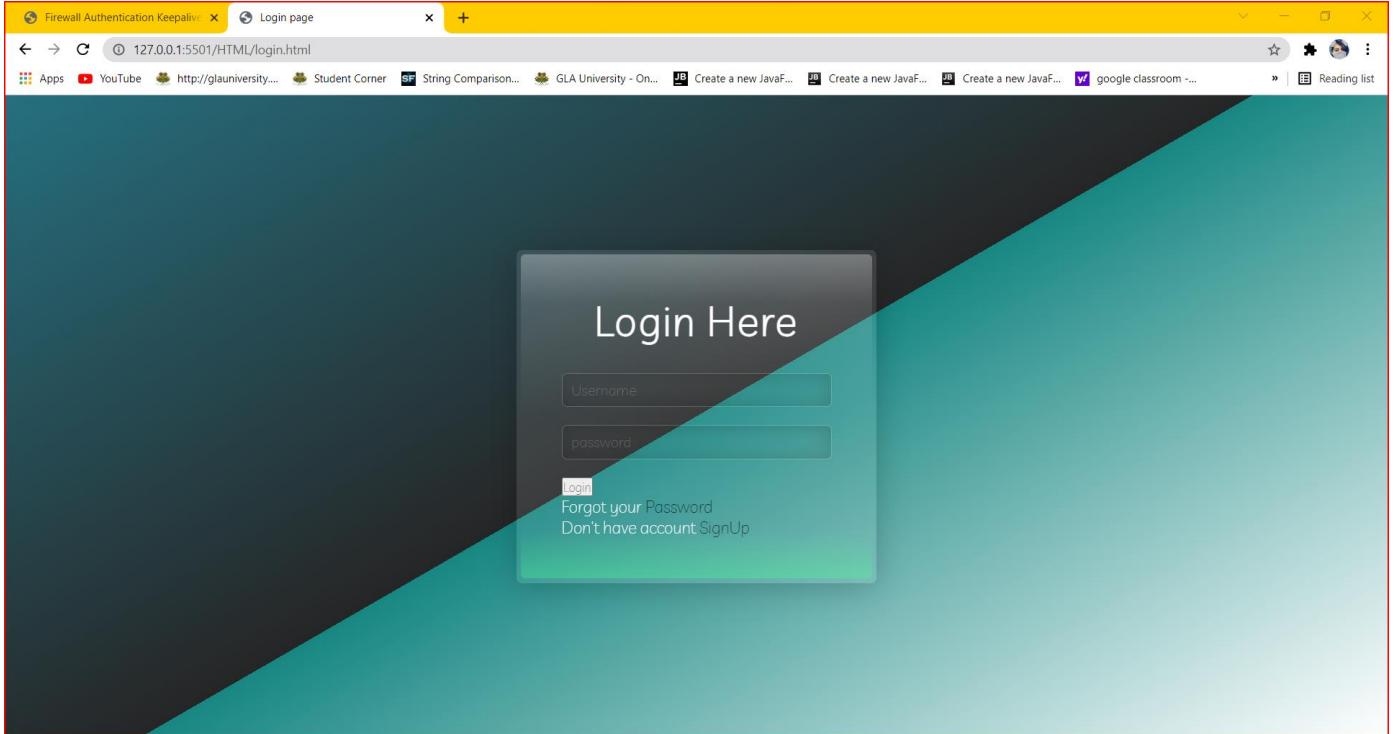
Email Address

Follow us

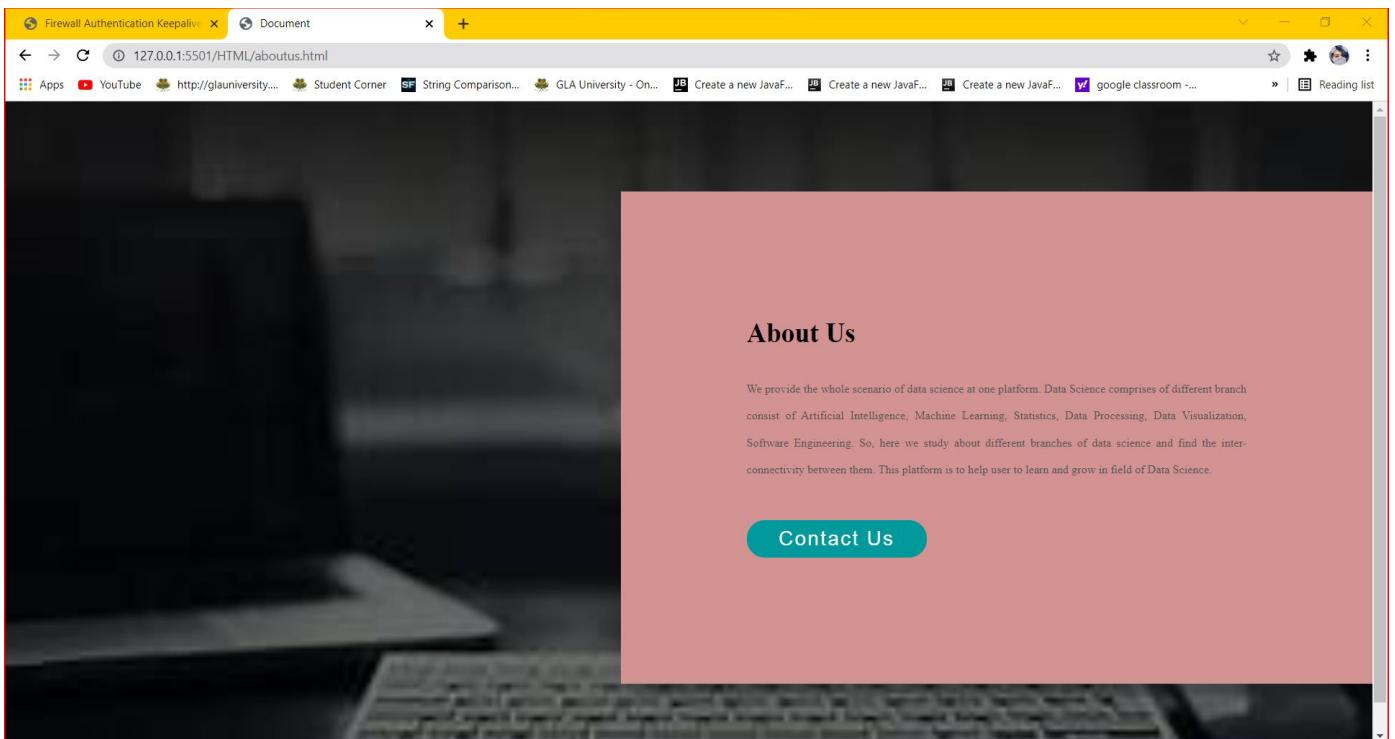
Let us be social



2. Login Page



3. About us Page



4. Contact us

Firewall Authentication Keepalive x contact us x +

127.0.0.1:5501/HTML/contact.html

Apps YouTube http://glauniversity.... Student Corner String Comparison... GLA University - On... JB Create a new JavaF... JB Create a new JavaF... JB Create a new JavaF... google classroom ... Reading list

The screenshot shows a contact form titled "Contact us" on a web browser. The background features a dark blue network graph with numerous nodes and connections. The form itself has a black background with white text and light blue input fields. On the left side, there are three sections with icons: a location pin for "Address", a phone receiver for "Phone", and an envelope for "E-mail". The "Address" section contains the text "Gla University ,Mathura". The "Phone" section contains the number "123-456-7890". The "E-mail" section contains the address "creativcosmos@gmail.com". To the right of these sections is a "Contact us" heading and a message "You can contact us using the details given below". Below this, there are four input fields: "Full name", "E-mail Id", and "Type your message...", each with a corresponding placeholder text. A pink "Submit" button is located at the bottom of the form. At the very bottom, there are social media icons for Facebook, Instagram, WhatsApp, and Twitter.

Contact us

You can contact us using the details given below

Address
Gla University ,Mathura

Phone
123-456-7890

E-mail
creativcosmos@gmail.com

Full name

E-mail Id

Type your message...

Submit

f i w t

5. Content of Artificial Intelligence

The screenshot shows a web browser window with the title bar "Artificial intelligence". The address bar displays the URL "127.0.0.1:5501/HTML/ai%20and%20ml.html". The page content is titled "—Artificial Intelligence—". Below the title, there is a section titled "What is artificial Intelligence-". It contains text about AI being a theory and development of computer systems that can perform tasks requiring human intelligence, with examples like speech recognition and decision-making. It also mentions AI as the simulation of human intelligence processes by machines. A list of specific applications follows: Expert systems, Natural language processing, Speech recognition, and Machine vision. Another section titled "Types Of Artificial Intelligence-" lists four types: Reactive, Limited memory, Theory of mind, and Self-aware. A final section titled "Need Of AI-" discusses how AI makes it possible for machines to learn from experience and perform human-like tasks, mentioning chess-playing computers and self-driving cars.

—Artificial Intelligence—

What is artificial Intelligence-

Artificial intelligence is a theory and development of computer systems that can perform tasks that normally require human intelligence. For Example - Speech recognition, decision-making, visual perception, are features of human intelligence that artificial intelligence may possess. Artificial intelligence is the simulation of human intelligence processes by machines, especially computer systems. Specific applications of AI include-

- Expert systems
- Natural language processing
- Speech recognition
- Machine vision

Types Of Artificial Intelligence-

- Reactive
- Limited memory
- Theory of mind
- Self-aware.

Need Of AI-

Artificial intelligence (AI) makes it possible for machines to learn from experience, adjust to new inputs and perform human-like tasks. Most AI examples that you hear about today – chess-playing computers to self-driving cars – rely heavily on deep learning and natural language processing.

The screenshot shows a Microsoft Edge browser window with a yellow header bar. The title bar says "Firewall Authentication Keepalive" and the tab bar says "Artificial intelligence". The address bar shows the URL "127.0.0.1:5501/HTML/ai%20and%20ml.html". The main content area displays the following text:

Fact related to Artificial Intelligence-

1. AI will Become Smarter than Humans-
AI can learn anything quickly, meaning its intelligence is increasing. In 2013, AI had the same intelligence as a 4 year old. By 2029, AI will have the same intelligence level as adult humans.

2. Humans can Develop Romantic Relationship with AI-
David Levy, a student in this field, believes that marriage between humans and robots will be made legal by 2050. At the moment robots are not so advanced enough for this.

3. AI can Write-
A robot wrote an article on an earthquake in California on the Los Angeles Times website, gathering data from a seismograph.

4. AI can Repair Itself-
A robot that rebuilt itself after noticing its performance had dropped after losing two of its six legs. The robot did not know what the problem was but fixed it by trial and error.

5. Most AI Bots are Female-
Studies show that most of the population prefer the sound of a female voice over a male voice. It is because if you ask voice assistants like Alexa, Siri a question, you will be answered by a pleasant and polite woman's voice.

Application of Artificial Intelligence

- AI in E-Commerce.
- AI in Navigation.
- AI in Navigation.
- AI in Human Resource.

6. Machine Learning content

—Machine Learning—

Introduction

A machine learning process begins by feeding the machine a lots of data, by using this data the machine is trained to detect hidden insights and trends. These insights are then used to build a Machine Learning model by using an algorithm in order to solve the problem. Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.

Data <-> Training the Machine <-> Building a Model <-> Predicting Outcome.

Machine Learning is a subset of Artificial Intelligence.

Why Machine Learning is important ?

- Increase in Data Generation.
- Improve Decision Making
- Solve Complex Problem
- Uncover patterns and trends in data

Application of Machine Learning-

- Automatic language translation.
- Speech Recognition.
- Traffic Prediction.

Types Of Machine learning-

- Supervised Machine Learning

```
graph TD; SL[Supervised Learning] --> C[Classification]; SL --> R[Regression]; C --> SVM[Support Vector Machine]; C --> RF[Random Forest]; C --> KNN[k-Nearest Neighbor]; C --> DT[Decision Tree]; R --> RT[Regression Tree]; R --> BM[Bayesian Model]; R --> RA[Regression Analysis]
```

- Unsupervised Machine Learning

Unsupervised Learning

```
graph TD; UL[Unsupervised Learning] -- Association --> UL; UL -- Clustering --> UL; UL -- Grouping --> UL
```

Supervised learning problems can be further grouped into regression and classification problems.

Classification: A classification problem is when the output variable is a category, such as "red" or "blue" or "disease" and "no disease".

Weaknesses: Unconstrained, individual trees are prone to overfitting. Logistic regression may underperform when there are multiple or non-linear decision boundaries. This method is not flexible, so it does not capture more complex relationships.

Strengths: Classification tree perform very well in practice and Outputs always have a probabilistic interpretation, and the algorithm can be regularized to avoid overfitting.

- **Regression:**

A regression problem is when the output variable is a real value, such as "dollars" or "weight". - Regression technique predicts a single output value using training data.

Some popular examples of supervised machine learning algorithms are:

Some popular examples of supervised machine learning algorithms are:

- Linear regression for regression problems.
- Random forest for classification and regression problems.
- Support vector machines for classification problems.

Supervised machine learning uses training data sets to achieve desired results. These data sets contain inputs and the correct output that helps the model to learn faster. For example, you want to train a machine to help you predict how long it will take you to drive home from your workplace.

Here, you start by creating a set of labeled data. This data includes:

- Weather conditions
- Time of the day
- Holidays

7. Statistics related to Data Science

—Statistics related to Data Science—

Statistics

The most important aspect of any Data Science approach is how the information is processed. Those possibilities in Data Science are known as Statistical Analysis. Most of us wonder how can data in the form of text, images, videos, and other highly unstructured formats get easily processed by Machine Learning models.

Use of Statistics in Data Science

- Statistical features is probably the most used statistics concept in data science. It's often the first stats technique you would apply when exploring a dataset and includes things like bias, variance, mean, median, percentiles, and many others. It's all fairly easy to understand and implement in code!
- Statistics is a wide concept limiting not just to what exists but what can be derived out of existing techniques to build something new. Hence, Statistics is very important for Data Science as it helps to understand existing solutions as well as digging out new developments.
- Statistics helps in providing a better understanding and accurate description of nature's phenomena.
- Statistics helps in collecting appropriate quantitative data.
- Statistics helps in the proper and efficient planning of a statistical inquiry in any field of study.

Terminology

- Probability

Probability is the basic need for understanding the possibilities. To start with let us take a very basic example – What are the chances that Team A is going to win the football match against Team B. To derive this answer, we might require 100 people to give their respective votes – Number of Samples. Based on those votes we can have a chance of which team can win the game.

Hypothesis Testing in R

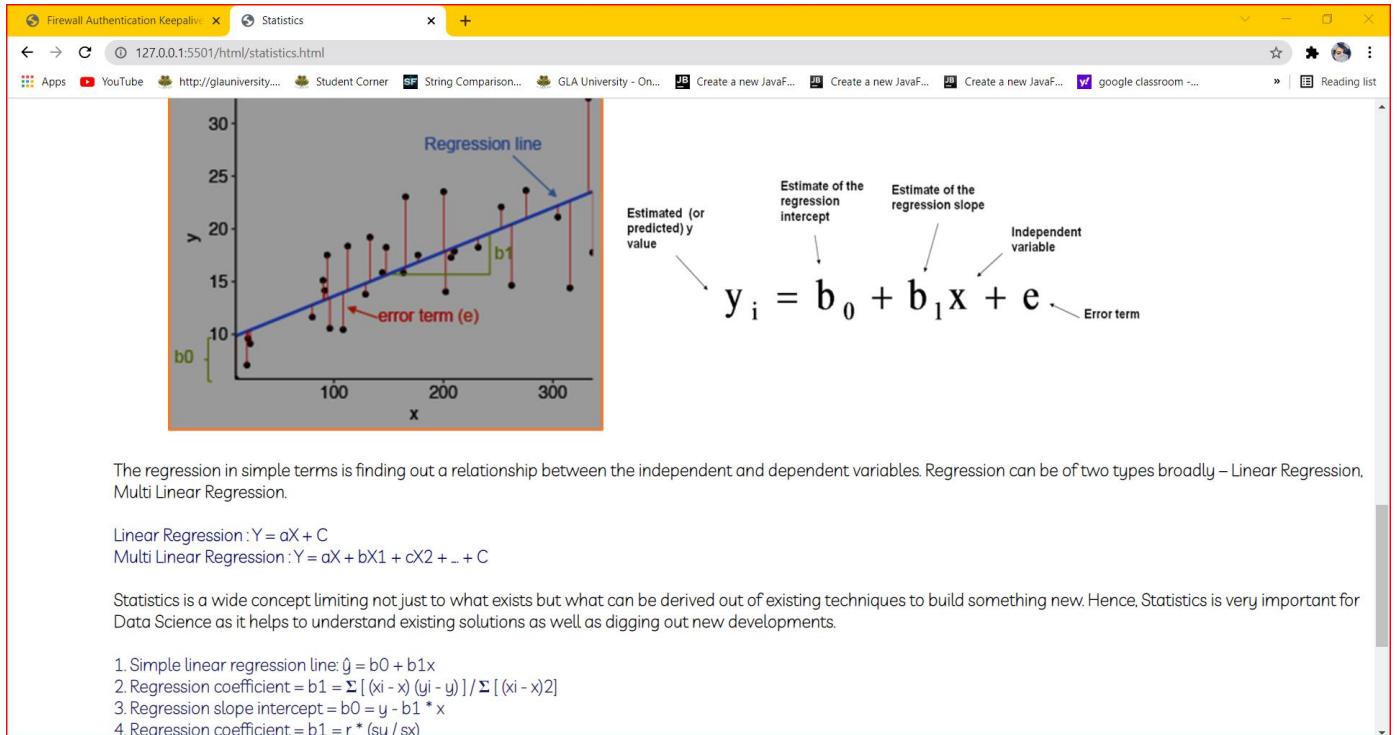
If the value is less than the threshold value, then we reject the alternative hypothesis and accept the null hypothesis.

If the values goes beyond the threshold value, then we accept the alternative hypothesis and reject the other one.

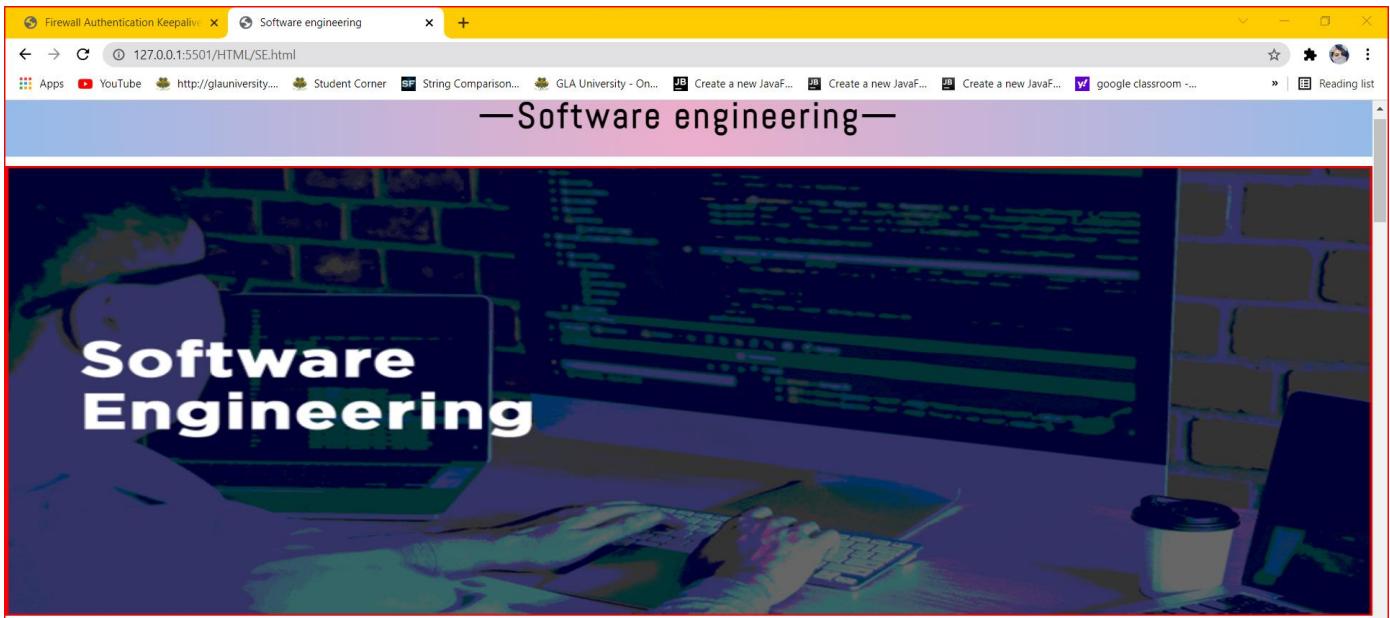
If we know whether to perform some action or not. Will those actions give a positive result or a negative result then we can have an added advantage of doing the right things. Hypotheses Testing gives identifying the situation where the action should be taken or not based on what results will it produce.

There are other tests as well like A/B Testing, Z Test, T-Test, Null Hypothesis with similar relevance.

1. Standardized test statistic = $(\text{Statistic} - \text{Parameter}) / (\text{Standard deviation of statistic})$
2. One-sample z-test for proportions: $z\text{-score} = z = (p - P_0) / \sqrt{p * q/n}$
3. Two-sample z-test for proportions: $z\text{-score} = z = [(p_1 - p_2) - d] / SE$
4. One-sample t-test for means: $t\text{-statistic} = t = (x - \mu) / SE$
5. Two-sample t-test for means: $t\text{-statistic} = t = [(x_1 - x_2) - d] / SE$
6. Matched-sample t-test for means: $t\text{-statistic} = t = [(x_1 - x_2) - D] / SE = (d - D) / SE$
7. Chi-square test statistic = $\chi^2 = \sum [(O \text{ - } E)^2 / E]$

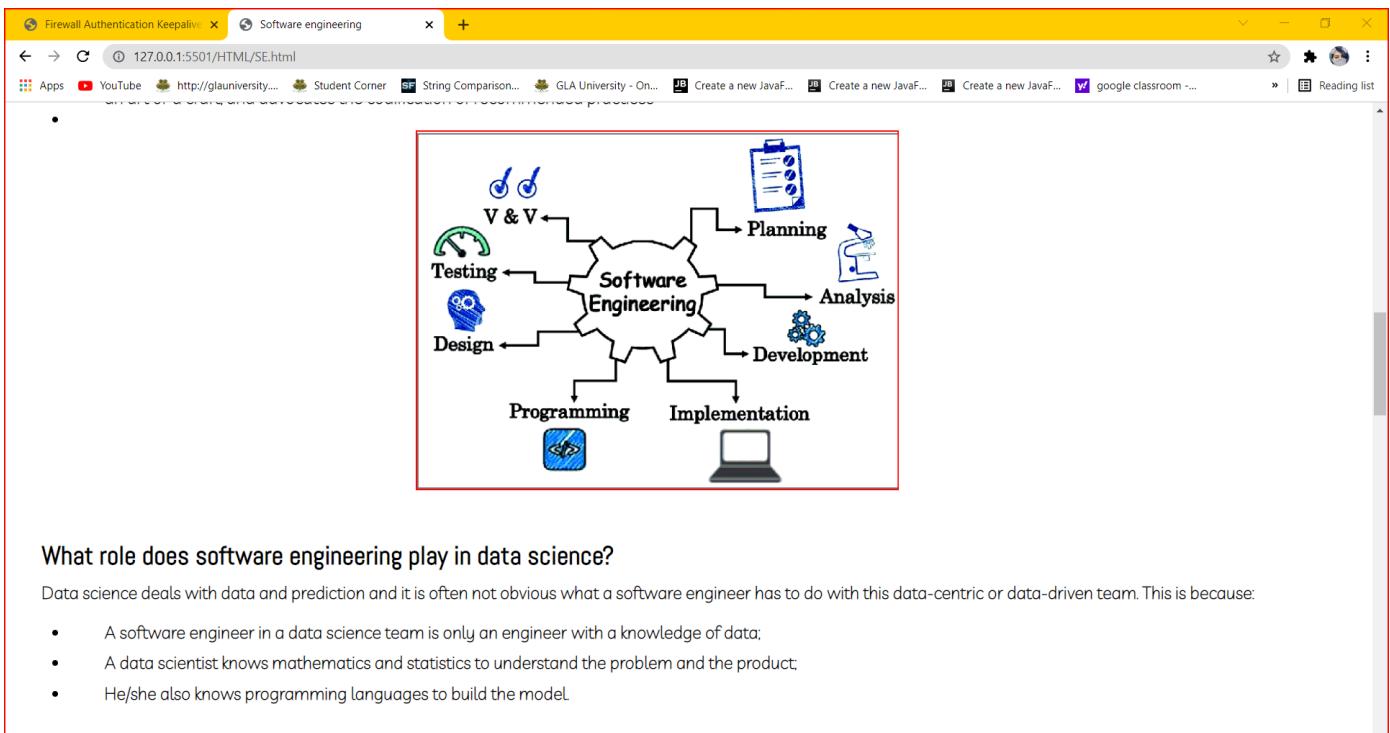


8.Content page of software engineering



What is Software Engineering

Software engineering is the systematic application of engineering approaches to the development of software. A software engineer is a person who applies the principles of software engineering to design, develop, maintain, test, and evaluate computer software. Modern processes use software versioning.



What role does software engineering play in data science?

Data science deals with data and prediction and it is often not obvious what a software engineer has to do with this data-centric or data-driven team. This is because:

- A software engineer in a data science team is only an engineer with a knowledge of data;
- A data scientist knows mathematics and statistics to understand the problem and the product;
- He/she also knows programming languages to build the model.

Responsibility Of The Software Engineer

A software engineer comes as a help broadly when the data is to be turned into a scalable product by adding extra hardware, enhancing the data's performance. In other words, his job is to productise the data science work so that the team can serve the external customers.

Building APIs:
His daily job involves creating Application Programming Interfaces (APIs), specifying how software components should interact and create a user interface. He/she has to ensure that the APIs created from the model is scalable, flexible and reliable.

Model examination:

The final product relies totally on the software engineer. He/she has to make sure that the model made by the data scientist can be used as a common model and that it can be easily managed. *By easy management, it means that he/she has to make sure that the model can be easily moderated to suit the other product requirements as well.*

Model testing and deploying:

Any model big or small, complex or easy, made by data scientists must be tested. His job is to review the code or the model created by the data scientist. Unit testing, branch testing, integration testing, security testing of the model is a part of his job. After testing, he/she takes a decision to deploy the model.

Skillset

The skills that they require is usually Hadoop, SQL, NoSQL, Hive, MapReduce, Pig. Some tools that are common to this profession are MySQL, MongoDB, DashDB and Cassandra, and it is expected that they be thorough in using them.



Outline of software engineering:

- **Software requirements:**
Requirements engineering is about the elicitation, analysis, specification, and validation of requirements for software. Software requirements can be of three different types. There are functional requirements, non-functional requirements, and domain requirements. The operation of the software should be performed and the proper output should be expected for the user to use. Non-functional requirements deal with issues like portability, security, maintainability, reliability, scalability, performance, reusability, and flexibility. They are classified into the following types: interference constraints, performance constraints (such as response time, security, storage space, etc), operating constraints, life cycle constraints (maintainability, portability, etc), and economic constraints.
- **Software design**
Software design is about the process of defining the architecture, components, interfaces, and other characteristics of a system or component. This is also called software architecture. Software design is divided into three different levels of design. The three levels are interface design, architectural design, and detailed design. *Interface design is the interaction between a system and its environment. This happens at a high level of abstraction along with the inner workings of the system.*
- **Software construction**
Software construction, the main activity of software development is the combination of programming, unit testing, integration testing, and debugging. Testing during this phase is generally performed by the programmer while the software is under construction, to verify what was just written and decide when the code is ready to be sent to the next step.
- **Software testing**
Software testing is an empirical, technical investigation conducted to provide stakeholders with information about the quality of the product or service under test, with different approaches such as unit testing and integration testing. It is one aspect of software quality. As a separate phase in software development, it is typically performed by quality assurance staff or a developer other than the one who wrote the code.
- **Software maintenance**
Software maintenance refers to the activities required to provide cost-effective support after shipping the software product. Software maintenance is modifying and updating software applications after distribution to correct faults and to improve its performance. Software has a lot to do with the real world and when the

9.Content page of data visualization

Firewall Authentication Keepalive Data Visualization 127.0.0.1:5501/HTML/dataVisualization.html Apps YouTube http://glauuniversity... Student Corner String Comparison... GLA University - On... Create a new JavaF... Create a new JavaF... Create a new JavaF... google classroom ~... Reading list

—Data Visualization—

What is Data Visualization ?

Data visualization is the practice of translating information into a visual context, such as a map or graph, to make data easier for the human brain to understand and pull insights from. The main goal of data visualization is to make it easier to identify patterns, trends and outliers in large data sets. The term is often used interchangeably with others, including information graphics, information visualization and statistical graphics.

Data visualization is the process of translating large data sets and metrics into charts, graphs and other visuals. The resulting visual representation of data makes it easier to identify and share real-time trends, outliers, and new insights about the information represented in the data.

Examples-->

- **Cinema:** Explaining a movie plot through data visualization

Inception is an American film, directed by Christopher Nolan and released in 2010, that focuses on the themes of dreams and reality. The movie's hero, Cobb, is an 'extractor', an agent that can enter someone's dreams and learn their secrets, who collaborates with others on industrial espionage missions.

- **Art:** Analyzing the color palettes of great artworks

Arthur Buxton has created a data visualization that shows an overview of the color palettes used by ten painters, including Monet, Gauguin, and Cézanne, over a period of ten years. These offer a new perspective on these artists, sorting them by the colors used rather than by art movement.

- **Philosophy:** A visual depiction of ideas

The data presented here is a direct reflection of concepts from everyday life. Ideas are presented and analyzed based on their significance, duration and the feelings that they evoke. Another interesting perspective on a common subject!

What Makes A Good

Firewall Authentication Keepalive Data Visualization 127.0.0.1:5501/HTML/dataVisualization.html Apps YouTube http://glauuniversity... Student Corner String Comparison... GLA University - On... Create a new JavaF... Create a new JavaF... Create a new JavaF... google classroom ~... Reading list

DATA VISUALIZATION

What Makes A Good

Data
- Find an interesting topic
- Look for an angle/purpose
- Collect and organize

Function
- Accurately represent data
- Make it clear and legible
- Make it purposeful and useful

Design
- Sketch out your process
- Create a thoughtful visual
- Find a balance with form and function

Why we need Data Visualization?

Data visualization gives us a clear idea of what the information means by giving it visual context through maps or graphs. This makes the data more natural for the human mind to comprehend and therefore makes it easier to identify trends, patterns, and outliers within large data sets.

In advanced analytics, data scientists are creating machine learning algorithms to better compile essential data into visualizations that are easier to understand and interpret.

Firewall Authentication Keepalive Data Visualization 127.0.0.1:5501/HTML/dataVisualization.html Apps YouTube Student Corner String Comparison... GLA University - On... Create a new JavaF... Create a new JavaF... Create a new JavaF... google classroom -... Reading list

Why we need Data Visualization?

Data visualization gives us a clear idea of what the information means by giving it visual context through maps or graphs. This makes the data more natural for the human mind to comprehend and therefore makes it easier to identify trends, patterns, and outliers within large data sets.

In advanced analytics, data scientists are creating machine learning algorithms to better compile essential data into visualizations that are easier to understand and interpret.

Specifically, data visualization uses visual data to communicate information in a manner that is universal, fast, and effective. This practice can help companies identify which areas need to be improved, which factors affect customer satisfaction and dissatisfaction, and what to do with specific products (where should they go and who should they be sold to). Visualized data gives stakeholders, business owners, and decision-makers a better prediction of sales volumes and future growth.

What Makes a Good Visualization?

The diagram illustrates the three pillars of a good visualization:

- story (concept)**: Includes research doc, script, article outline, schematic, wireframe, plot, and storyboard.
- goal (function)**: Includes proof of concept, prototype, template, detailed sketch, and rough sketch.
- visual form (metaphor)**: Includes art and metaphor.

What Is Good Data Visualization?

- Show the data and have a clear purpose
- Make the graphic elements background, so the substance is what the viewer is thinking about
- Never distort the data
- Present many numbers in a small space
- Make large data coherent
- Think of the viewer and what they're looking at, help guide their eye to what's important
- Go from big picture to more detailed
- Integrate statistical and verbal description of data
- Have the graphic reveal the data

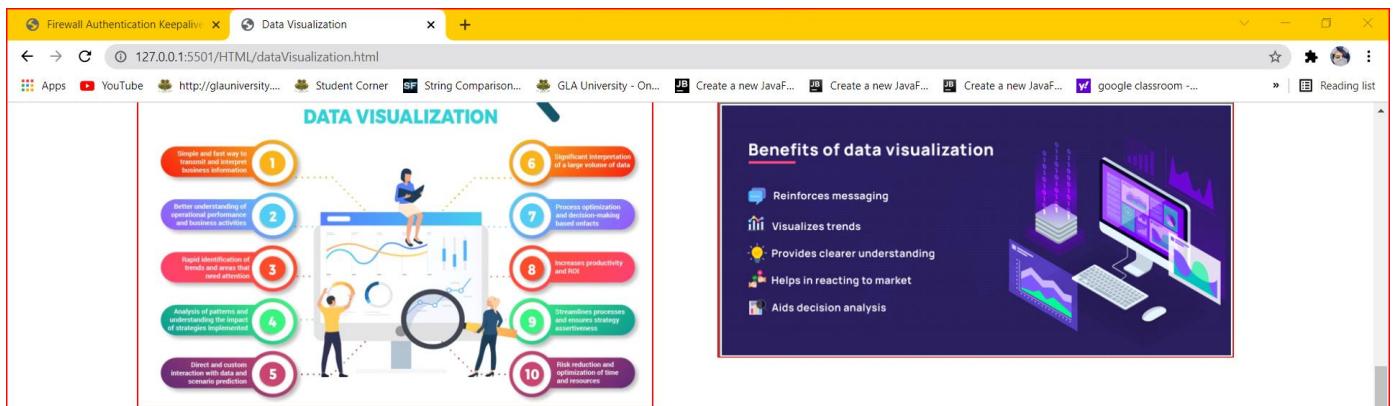
Firewall Authentication Keepalive Data Visualization 127.0.0.1:5501/HTML/dataVisualization.html Apps YouTube Student Corner String Comparison... GLA University - On... Create a new JavaF... Create a new JavaF... Create a new JavaF... google classroom -... Reading list

Benefits of Data Visualization?

Data visualization positively affects an organization's decision-making process with interactive visual representations of data. Businesses can now recognize patterns more quickly because they can interpret data in graphical or pictorial forms.

Here are some more specific ways that data visualization can benefit an organization:

- Correlations in Relationships:**
Without data visualization, it is challenging to identify the correlations between the relationship of independent variables. By making sense of those independent variables, we can make better business decisions.
- Trends Over Time:**
While this seems like an obvious use of data visualization, it is also one of the most valuable applications. It's impossible to make predictions without having the necessary information from the past and present. Trends over time tell us where we were and where we can potentially go.
- Frequency:**
Closely related to trends over time is frequency. By examining the rate, or how often, customers purchase and when they buy gives us a better feel for how potential new customers might act and react to different marketing and customer acquisition strategies.
- Examining the Market:**
Data visualization takes the information from different markets to give you insights into which audiences to focus your attention on and which ones to stay away from. We get a clearer picture of the opportunities within those markets by displaying this data on various charts and graphs.
- Risk and Reward:**
Looking at value and risk metrics requires expertise because, without data visualization, we must interpret complicated spreadsheets and numbers. Once information is visualized, we can then pinpoint areas that may or may not require action.
- Reacting to the Market:**
The ability to obtain information quickly and easily with data displayed clearly on a functional dashboard allows businesses to act and respond to findings swiftly and helps to avoid making mistakes.



Tools used in data visualization:

- **Tableau**

Tableau is a data visualization tool that can be used by data analysts, scientists, statisticians, etc. to visualize the data and get a clear opinion based on the data analysis. Tableau is very famous as it can take in data and produce the required data visualization output in a very short time. And it can do this while providing the highest level of security with a guarantee to handle security issues as soon as they arise or are found by users.

Tableau also allows its users to prepare, clean and format their data and then create data visualizations to obtain actionable insights that can be shared with other users. Tableau is available for the individual data analyst or at scale for business teams and organizations. It provides a 14-day free trial followed by the paid version.

- **Looker**

Looker is a Looker data visualization tool that can go in-depth in the data and analyze it to obtain useful insights. It provides real-time dashboards of the data for more in-depth analysis so that businesses can make instant decisions based on the data visualizations obtained. Looker also provides connections with Redshift,

Tools used in data visualization:

- **Tableau**

Tableau is a data visualization tool that can be used by data analysts, scientists, statisticians, etc. to visualize the data and get a clear opinion based on the data analysis. Tableau is very famous as it can take in data and produce the required data visualization output in a very short time. And it can do this while providing the highest level of security with a guarantee to handle security issues as soon as they arise or are found by users.

Tableau also allows its users to prepare, clean and format their data and then create data visualizations to obtain actionable insights that can be shared with other users. Tableau is available for the individual data analyst or at scale for business teams and organizations. It provides a 14-day free trial followed by the paid version.

- **Looker**

Looker is a Looker data visualization tool that can go in-depth in the data and analyze it to obtain useful insights. It provides real-time dashboards of the data for more in-depth analysis so that businesses can make instant decisions based on the data visualizations obtained. Looker also provides connections with Redshift, Snowflake, BigQuery, as well as more than 50 SQL supported dialects so you can connect to multiple databases without any issues.

Looker data visualizations can be shared with anyone using any particular tool. Also, you can export these files in any format immediately. It also provides customer support wherein you can ask any question and it shall be answered. A price quote can be obtained by submitting a form.

- **Zoho Analytics**

Zoho Analytics is a Business Intelligence and Data Analytics software that can help you create wonderful looking data visualizations based on your data in a few minutes. You can obtain data from multiple sources and mesh it together to create multidimensional data visualizations that allow you to view your business data across departments. In case you have any questions, you can use Zia which is a smart assistant created using artificial intelligence, machine learning, and natural language processing.

Zoho Analytics allows you to share or publish your reports with your colleagues and add comments or engage in conversations as required. You can export Zoho Analytics files in any format such as Spreadsheet, MS Word, Excel, PPT, PDF, etc. The pricing options available for this software include a basic plan with A\$34.1/month billed yearly.

- **Sisense**

Sisense is a business intelligence-based data visualization system and it provides various tools that allow data analysts to simplify complex data and obtain insights for their organization and outsiders. Sisense believes that eventually, every company will be a data-driven company and every product will be related to data in some way. Therefore, it tries to provide various data analysis tools to business teams and data analysts so that they can help make their

Firewall Authentication Keepalive Data Visualization +

127.0.0.1:5501/HTML/dataVisualization.html

Apps YouTube http://glauniversity.... Student Corner String Comparison... GLA University - On... Create a new JavaF... Create a new JavaF... Create a new JavaF... google classroom -... Reading list

Real Time Application Data Visualization:

Real-time data visualization takes visuals to the next level by letting you update charts and graphs in real-time. Having real-time data available helps stakeholders to make better decisions that are based on actual data rather than on intuition. Data visualisation tools make it easy to view and comprehend trends, outliers, and patterns in data by utilising visual components like charts, graphs, and maps.

How do we build real-time visualizations?

At its core, a real-time visualization is made up of two components: a stream of data and a user interface.

- Stream of Data**
A stream of data is a constant flow of data where data packets are sent whenever that data is updated. This could be weather forecasts, stock prices, revenue, or Tweets, pretty much any data group that changes over time. Because a stream of data is active, it needs to have persistence, an open connection that allows data to be transferred freely.
- User Interface**
A real-time visualization requires a flexible, interactive user interface where changes are easily displayed based on changes sent from the stream of data.



SOFTWARE TESTING

Once source code has been generated, software must be tested to uncover as many errors as possible before delivery. It is very important to work the system successfully and achieve high quality of software. Testing includes designing a series of test cases that have a high likelihood of finding errors by applying software-testing techniques. System testing makes logical assumptions that if all the parts of the system are correct, the goal will be successfully achieved. The system should be checked logically. Validations and cross checks should be there. Avoid duplications of record that cause redundancy of data. In other Words, Testing is the process of evaluating a system or its component(s) with the intent to find whether it satisfies the specified requirements or not. It is executing a system

in order to identify any gaps, errors, or missing requirements in contrary to the actual requirements.

The preliminary goal of implementation is to write source code and internal documentation so that conformance of the code to its specifications can be easily verified, and so that debugging, testing and modifications are eased. This goal can be achieved by making the source code as clear and straightforward as possible. Simplicity, clarity and elegance are the hallmark of good programs, obscurity, cleverness, and complexity are indications of inadequate design and misdirected thinking. Source code clarity is enhanced by structured coding techniques, by good coding style, by, appropriate supporting documents, by good internal comments, and by feature provided in modern programming languages. The implementation team should be provided with a well-defined set of software requirement, an architectural design specification, and a detailed design description. Each team member must understand the objectives of implementation.

4.1 TERMINOLOGY

Error The term error is used in two ways. It refers to the difference between the actual output of software and the correct output, in this interpretation, error is essential a measure of the difference between actual and ideal. Error is also used to refer to human action that result in software containing a defect or fault.

Fault is a condition that causes to fail in performing its required function. A fault is a basic reason for software malfunction and is synonymous with the commonly used term Bug.

Failure is the inability of a system or component to perform a required function according to its specifications. A software failure occurs if the behavior of the software is different from the specified behavior. Failure may be caused due to functional or performance reasons.

4.2 TYPES OF TESTING

a. Unit Testing The term unit testing comprises the sets of tests performed by an individual programmer prior to integration of the unit into a larger system. A program unit is usually small enough that the programmer who developed it can test it in great detail, and certainly in greater detail than will be possible when the unit is integrated into an evolving software product. In the unit testing the programs are tested separately, independent of each other. Since the check is done at the program level, it is also called program teasing.

b. Module Testing A module and encapsulates related component. So can be tested without other system module.

c. Subsystem Testing Subsystem testing may be independently design and implemented common problems are sub-system interface mistake in this checking we concenton it. There are four categories of tests that a programmer will typically perform on a program unit.

i Functional test

ii Performance test

iii Stress test

iv Structure test

Functional Test Functional test cases involve exercising the code with Nominal input values for which expected results are known; as well as boundary values (minimum values, maximum values and values on and just outside the functional boundaries) and special values.

Performance Test Performance testing determines the amount of execution time spent in various parts of the unit, program throughput, response time, and device utilization by the program unit. A certain amount of avoid expending too much effort on fine-tuning of a program unit that contributes little to the overall performance of the entire system. Performance testing is most productive at the subsystem and system levels.

Stress Test Stress test are those designed to intentionally break the unit. A great deal can be learned about the strengths and limitations of a program by examining the manner in which a program unit breaks.

Structure Test Structure tests are concerned with exercising the internal logic of a program and traversing particular execution paths. Some authors refer collectively to functional performance and stress testing as “black box” testing. While structure testing is referred to as “white box” or “glass box” testing. The major activities in structural testing are deciding which path to exercise, deriving test date to exercise those paths, determining the test coverage criterion to be used, executing the test, and measuring the test coverage achieved when the test cases are exercised.

Conclusion

- We have completed our project within time limit with the coordination of our team members under the supervision of our mentor Ms. Ruchi Gupta.
- Our project repository is available at
<https://github.com/ishasingh23/Erudition-in-Data-Science-and-Machine-learning>
- Our project live website link is
<https://ishasingh23.github.io/Erudition-in-Data-Science-and-Machine-learning/>

Bibliography

- www.google.com
- www.geeksforgeeks.org
- www.youtube.com
- www.w3schools.com
- www.beta-labs.in

