Big Data and Cloud Computing Exam

Big data is defined as data with greater diversity, coming in higher volumes and with greater velocity. This is sometimes referred to as the three V's. While big data advances year after year, there are currently more than five V's, with more being developed. Simply said, big data refers to larger, more complicated data volumes, particularly from new data sources. These data sets are so large that typical data processing technologies just cannot handle them. However, these huge amounts of data may be used to solve business challenges that were previously unsolvable. In this report we will go in depth creating a company, The Moon Security to show the importance and differences within cloud computing and big data.

Task 1 –

Cloud computing also commonly referred to as 'The Cloud,' is one of the most efficient mechanisms that delivers computing services (Vennam, 2020). On the contrary from Cloud computing there are traditional computing. As its name suggest, traditional computing consists of storing digital assets in physical datacentres and managing a network for daily operations through a complete physical network infrastructure (GeeksforGeeks, 2021). Migration to the cloud is the process of moving data and applications from on-site servers to those provided by public cloud providers such as Amazon and Microsoft, but also between different clouds (EDUCBA, 2020).

These cloud computing services are efficient and viable for a wide range of users due to variety of services it offers, as it allows and gives them freedom of choice of the time, place, and extent of when, where, and how the company wants to adopt it. These services are divided with their own benefit and features that are suitable for the companies requites and goals. (Tambe, 2019). The services include Infrastructure-as-a-service (Iaas), platform-as-a-service (Paas) & software-as-a-service (Saas).

In IaaS, the resources such as Virtualisation, Networking, servers, and storage that are provided by a third party are supplied by bare computers. By doing so, the company, in this case, has a full overview of its infrastructure management and can make responsible changes to it (Boisvert, et al., 2020). Since it is accessible via the internet and is a virtual asset, it is quicker, easier, and more cost-effective than traditional approaches. Applications are installed and run through IaaS, while enterprise administrators manage and maintain software layers

installed by their organizations. In addition to its ability to adapt to changes in performance and cost, IaaS provides the company's systems with the flexibility to meet changing demands for applications and systems. The increase in scalability gives the company an overall expansion in the business while maintaining the quality of its products and services.

Cost optimization is a disciplined component for companies from a standardized standpoint (Azure Microsoft, 2022). Where the emphasis is on company budgeting and financial reduction. Companies employ a variety of resources, devices, and other methods to improve their operations and goods. There are additional fees and costs associated with using these resources. Giving that the customers of the IaaS model pay for their usage on a per-user basis, usually per hour, week, or month. Customers are also charged based on the amount of virtual machine space they use by some IaaS providers. With the pay-as-you-go model, in-house hardware, and software it can be deployed without incurring capital expenditures. (Boisvert, et al., 2020). As a result, the enterprise can deploy and further commit to this model for a long time or temporarily as it relates to various projects.

PaaS, on the other hand, is based on the infrastructure model, with providers extending it by hosting operating systems, middleware, other framework servers, and other services on physical infrastructure (Liquid State, 2019). Hardware, installation, confederation, and upgrades are all managed by providers. It provides middleware and other run-time services. As a result of this, businesses may place a greater emphasis on their own applications and data, reducing their infrastructure management responsibilities. It let businesses to deploy apps and data without having to worry about the underlying architecture.

PaaS enables quick development with faster and more frequent release of functionality, which directly supports business agility. (Wikipedia Contributors, 2022)As a result, customers will experience improved responsiveness, reduced performance problems and more innovative features, as well as improved operational efficiency. By implanting processes that improve employee workflow, both employees, customers, and the company itself will benefit of this matter (Meegan, 2016).

Most cloud platforms have a worldwide presence, allowing businesses to optimize performance by deploying workloads closer to their users or customers. In SaaS applications, multitenant architecture is employed to separate client data. Software upgrades, bug fixes, and other regular program maintenance are handled by the SaaS provider, while users interact

with the software through a web browser (IBM, 2020). It is completely administered and maintained by the providers using a web browser, in other words, it is web hosted software. It eliminates the need for infrastructure, platforms, and on-premises software purchases and maintenance.

Moon security being an international company that markets smart homes, sensors, IoT devices, alarms, CCTV cameras and etc. A cloud migration would be beneficial in terms of the relationship between IoT and cloud computing. In general, IoT is most prevalent in manufacturing, transportation, and utility firms, which deploy sensors and other IoT devices; but it has also found use cases in the agricultural, infrastructure, and home automation industries, propelling certain organizations towards digital transformation (McKenna Consultants, 2021).

TASK 2 –

91% of businesses used a public cloud, and 72% used a private one, as reported by Right Scale's 2019 State of the Cloud Report. Many enterprises use both options - with 69% choosing hybrid cloud solutions (WebTribunal, 2022). A company may select a complex migration strategy from the start due to special business requirements and the nature of the assets to be migrated. That is why the companies need to make a comprehensive approach to understand what to move and make correct planning. That is why there are many ways to migrate. (Citrix.com, 2021)

If we happened to find a new solution for companies that have sensitive, big data and only want to be part of cloud providers, we can come up with something called Re-assign. This system is made up of several self-driving computers that communicate with one another via a network. They each have their own memory. According to Outposts, there are a few risks that we should be aware of before we implement the process of migrating. One of the first risks the website is describing is the *Unclear Cloud Migration Strategy*. Not everything can be migrated. It is not recommended to store financial records, client information, and other sensitive data, in the public cloud due to GDPR (Vishnyov, 2021). Therefore, most companies consider the hybrid strategy - combining a public cloud and an on-premises data center, to host sensitive data at their on-premises data centers and utilize cloud providers for their benefits. As a customer, you will manage to work with the cloud providers once to give them the planning of your infrastructure. After that, all computers in the system communicate

with each other to accomplish a common goal. It may be said that the cloud providers and the company on each computer have different requirements. Therefore, the system will coordinate resources (or facilitate communication with other nodes). To fulfill those requirements. The communication will either be accomplished via messages or over the phone. Furthermore, a single major problem can be broken down into tasks, and each of those is computed by a separate computer in the system. And then, it will either connect to Rehost or repurchase, which is the 6 Rs in the migration strategy. Rehost is the “lift and shift”, Transferring your existing applications and systems exactly as they are, while the cloud provider will tailor SaaS to your specific business needs when you repurchase. Repurchasing is a great and fast way to get access to cloud-based SaaS (Beacon Telecom, 2018).

There are several products or services offered through IaaS and PaaS. Amazon EC2, Microsoft Azure, and Google Compute Engine provide general-purpose virtual machines (VMs). All three providers offer multiple versions of Linux and Windows. (A Cloud Guru, 2022). Amazon E3, Google Cloud Storage, and Microsoft Azure Blob Storage are storage products. They all provide versioning, encryption at rest, fine-grained security, and the ability to make an object public or completely private. Providing the reason why it is important for customers to have it as a part of their infrastructure. The more performant and redundant the storage class, the higher the price. They also provide other types of products or options to reduce costs for less frequently accessed data. Amazon S3 infrequent access, Azure blob storage cool tier, and GCP nearline storage and oldline storage. Amazon elastic block store, Google persistent disks, and Solidate disks are cloud providers' block storage options. Block storage is what allows our virtual machines to run their operating systems. The same features are available on call, so you can use a fast SSD device or a traditional magnetic hard drive for less money. Block storage is what allows our virtual machines to run their operating systems. The same features are available on call, so you can choose to use a fast SSD device or a traditional magnetic hard drive for less money. You can use snapshots to backup, restore, and duplicate your block storage devices. This can be a good infrastructure for most organizations to follow, just in case something goes wrong with the data.

We can combine products from different cloud providers, especially the products above due to their IaaS and PaaS services, to create layers of IT resources that can be added to each other. Using the PaaS environment, organizations can create their own cloud services that can be made available to customers by adding them to the preceding layered architecture. As a customer, you will have access to the cloud architecture for a limited amount. It depends on

what the services offer. All the products above can make an infrastructure that gives the customer all access by their own decision, only choosing one service and getting all access to everything at the same time with the cloud providers. They provide customers with an open version of the products and have a closed version of the products, which they will later combine with the customer's editing to create a final product. It is important that the VMs is untouched by the customers due to errors that will be caused.

The products provide various types of advantages for the company and for their customers. Scalable computing power, storage space, and developing, maintaining, and running home services, as well as remote access to home devices, are a few advantages for the company (Domb, 2019). One of the big bonuses the company has is that it will give the workers from Singapore, Frankfurt, and any other places it will establish eventually, access to the products anytime and anywhere no matter which time of the day. But somehow, the company will be dealing with a lack of support. It will be a problem if the company is unable to access customer data. However, the biggest disadvantage for the company will be that all these products must be accessed via a network. According to Kyle Turco from Technology Advice, cloud providers reduce the necessary amount of hardware. Hardware upgrades can be expensive, inconvenient, and time-consuming for companies that are growing rapidly. Due to the ease of acquiring resources through cloud computing, these issues are alleviated. Moreover, repair or replacement costs of equipment are passed on to vendors (Turco, 2021). The storage products save space and will cut internal power costs because a company with one million customers will take an unnecessary amount of space and produce an unnecessary amount of heat.

Security is one of every business biggest concerns. It is a legal requirement for businesses to safeguard customer information. Authentication, access control, and encryption are among the security measures that cloud storage providers employ to protect their platforms and the data they process. Cloud providers storage products has also unlimited capacity. There is essentially no limit to how much data can be stored in the cloud, depending on the frequency of access, the accessibility, and the availability of the data (Globaldots, 2018). But the prices will be higher. The company can store, process, retrieve and recover their work anytime. While the businesses in the digitalization in any devices will somehow experience data loss. Products that have been used to build the infrastructure especially storage will enable rapid data recovery and help with loss prevention in a variety of emergency situations.

But how secure is business data in the cloud? There are major concerns in data security for companies when they switch to cloud providers. Customers' and cloud providers' data is private and sensitive. According to @Vanigupta20024, from GeekforGeeks blog page, “7 Most common Cloud computing challenges”, it is the company's responsibility to conduct authentication, identity management, data encryption, and access control even if the cloud service provider assures data integrity. Data breaches, malware infections, and identity theft are some of the security threats in the cloud that decrease the trust among your users. As a result, revenue and reputation can be adversely affected. (GeeksforGeeks, 2021). It will be impossible for the company to be fully responsible to protect customers' sensitive data while only settling on cloud providers. That is why it will fully be secured to follow multi-cloud strategy and hybrid cloud. According to MindInventory, The RightScale report revealed that 84% of companies are following a multi-cloud strategy and 58% have already implemented hybrid cloud technologies that combine public and private cloud technologies (Patel, 2020).

Task 3 –

Some big data challenges are problems with data growth, data security, and poor misunderstanding of big data. According to Chandan Gaur, a major challenge with massive data is storing these large sets of knowledge effectively. The amount of knowledge being stored in data centres, and databases is increasing rapidly. With time, these data sets grow exponentially, making it increasingly difficult to handle. The same goes for data growth. Data security is often pushed to the back burner as companies focus on understanding, storing, and analysing their data sets.

In most cases, professionals have not kept up with the rapid development of data management technologies. Businesses demand skilled data experts to deal with this new technology and big data tools. Other employees, on the other hand, may lack essential knowledge of the subject, such as what data is? How crucial it is? etc. Such Big Data Tools are frequently used by professionals who are not data science experts but have a basic understanding of the subject (Xenonstack, 2020).

To manage the growth of the business, it is important to handle these 4V’s. Why? Gartner's website defines big data as high-volume, high-velocity, and/or high-variety information assets that often require cost-effective, innovative forms of information processing to generate more insight, improve decision-making, and automate processes (Gartner, 2012). 4Vs will let the big data process effective and let the business grow without limits and any issues. It is unlikely that major problems will arise during big data processing.

Various framework, such as Apache Cassandra, Hadoop, and Spark can handle the problem due to their work like storing and processing. Hadoop, for example, uses a distributed file system to store big data. Customers with large files will break them down into smaller chunks and place them on various machines. They can also make copies that will go to different nodes when they break the file. In this way, even if one machine fails, the data is still safe on another. The map-reduce technique is used for processing big data. Thus, the processing becomes easier and faster. This data can then be analysed for a variety of purposes (Simplilearn, 2019).

Securing these vast volumes of knowledge is one of the most difficult aspects of big data. Companies are typically so busy with understanding, preserving, and analysing their data sets that they put data security off till later. This is not always a reasonable suggestion because unprotected data repositories might serve as a breeding ground for malicious hackers as mentioned earlier. When a company is exposed to malicious acts such as data breach, the company may be held liable. Although the company was the victim of a crime, it could still be held responsible for the circumstance (HG.org, 2022). This is due to the company's ability to keep the information secure.

An option can be to invest heavily in hiring experienced specialists. To get the most out of present employees, the company, as previously stated, must give training programs. Moon security should also invest in artificial intelligence/machine learning-powered knowledge analytics solutions.

Another problem to address is accessibility within the company’s data. Who has access to which data and at what degree of permission? And who has the authority to change them?

These obstacles must be resolved in a face-to-face dialogue between IT and end users to identify who should be permitted access. It is potential that the company should utilize a variant of the Bell-LaPadula model to implement access control; this phenomenon is used to assure security confidentially (Wikipedia Contributors, 2021). Furthermore, with so much big data coming in, security breach can occur at any stage of data processing. The security risk and potential solution have been provided at each stage, from data collecting through data storage, analysis, and processing.

Retailers can use big data to their advantage as well as a challenge. It is altering the company by allowing for the precise prediction of demand for tailored items. It may also provide

manufacturers additional lead time by identifying changes in client behaviour, allowing them to develop customized items virtually as effectively as those produced on a larger scale. In other words, it will let them efficiently and effectively plan their product assortment mix at the hyper-local level (Reliable Software Inc, 2019) . Organizing a large volume of data, extracting useful information from it, and using that information to enhance business is what big data is about, with help from cloud providers' products (Chan, 2018). Another advantage is that it can be used by any type of business and in a variety of ways, such as using sensors to evaluate equipment performance, improving distribution networks, better monitoring employee performance, and even sourcing top talent.

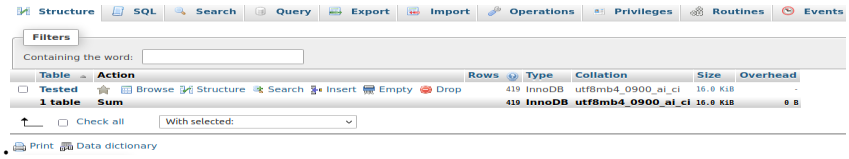
Practical Task

The "Titanic dataset" from Kaggle in CSV format is the dataset we used. We picked this dataset since it is the most famous historical accident of all time. This dataset details who was killed or injured in the accident. Passenger numbers, survivals, which class the passengers belong to, name, sex, age, number of siblings, parch, ticket, and fee are the data with which we were provided. The database "Titanic Dataset" has been used three times for similar purposes, according to Google Scholar. Two of these papers use the same "Titanic Dataset" to develop and teach machine learning approaches. The British University in Egypt uses one, while the Gangwon Embedded Software Cooperative Research Center at Gangneung-Wonju National University in South Korea uses the other. The other research is an article on predicting the Titanic's survivor rate. The "Titanic Dataset" is an extremely well-organized and prepared dataset, with very few codes and just one CSV-file containing crucial data, which may explain why several diverse studies used the same database.

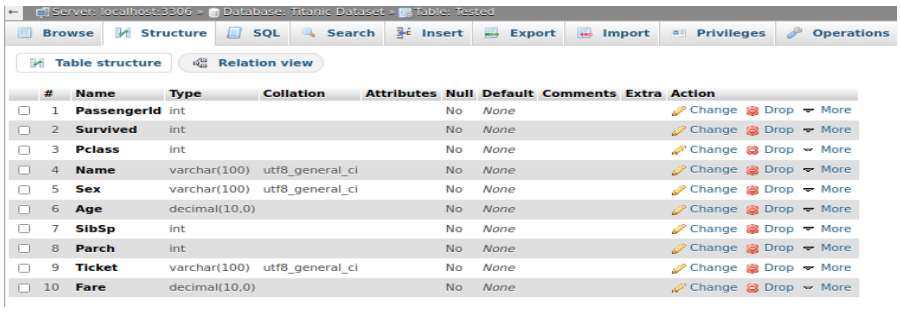
Task 1 –

We had to use a VM machine with Ubuntu to build a MySQL database on PhpMyAdmin for the practical exercise. We had to first prepare the VM machine by downloading MySQL-server, Apache2, PhpMyAdmin, and many more tools in order to activate and operate PhpMyAdmin. The remaining of the work was straightforward after preparing and starting PhpMyAdmin. We were able to build a database on phpMyAdmin by selecting New, Create Database, and specifying the desired database name. We named our database Titanic Dataset. Following these processes, tables were created. Structure was used to generate a suitable

table for our dataset. To make it easy to import, we named our table "Tested" after our excel file



After that, we carefully included each attribute in our excel file with the right data type and keys, making sure they all had the same provided name. In our case, the majority of data types were integers, with the exception of name, ticket, and sex, which were Varchar(65). Since age and fare involve decimals, the data type used is Decimal. As there were no columns that were related to one another in our dataset, we assigned the no key to all of them to signal that they are distinct. We could have used UNIQE, but there was no discernible difference , so we selected non.



The following step was to import our CSV– file into our database. Before we could do so, we needed to make some adjustments to our excel file. We needed to make a modification and ensure that it was saved as a CSV(MS-DOS)(*.csv). We used the IMPORT tool to import our data and saved it as a CSV file. The next step was to load the file into our database. After checking and doing so, we were able to import our dataset and get the data in our tables. There were 418 rows of data in our table.

Max – Min – Average values – Number of unique per feature.

To find Max, Min, Average and unique values we used similar code, changing the command.

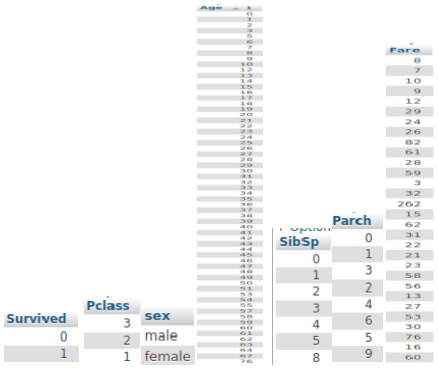
Max(Survived)	MAX(Pclass)	Max(Sex)	Max(Age)	Max(SibSp)	Max(PArch)	Max(Fare)
1	3	male	76	8	9	512

MIN(Survived)	MIN(Pclass)	MIN(Sex)	MIN(Age)	MIN(SibSp)	MIN(Parch)	MIN(Fare)
0	1	female	0	0	0	0

AVG(Survived)	AVG(Pclass)	AVG(Age)	AVG(SibSp)	AVG(PArch)	AVG(Fare)
0.3636	2.2656	24.0622	0.4474	0.3923	35.6268

Unique Data Storage

Our Dataset used 80 KiB storage space on our database



Space usage	
Data	80.0 KiB
Index	0 B
Overhead	
Effective	80.0 KiB
Total	80.0 KiB

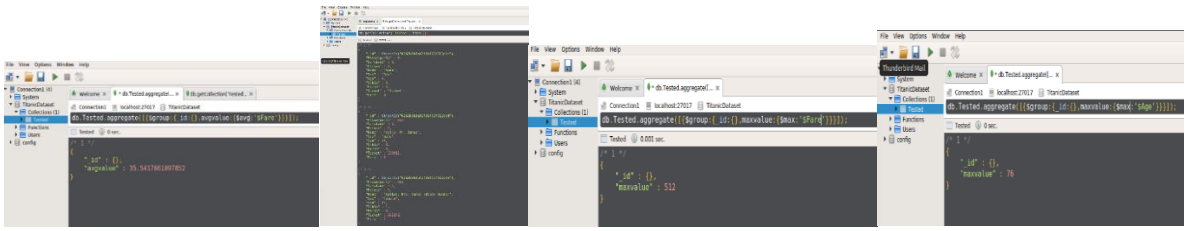
Task 2 –

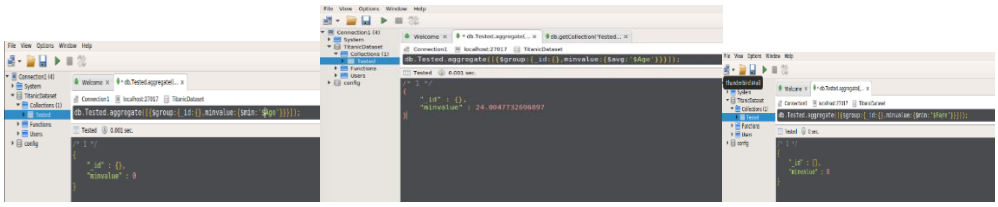
The next task was to import the MySQL database to a NoSQL database like MongoDB. The importing from MySQL to MongoDB is completed in four phases. Reviewing the data schemas is the first step. The second step is to export MySQL data. The third and the last step is to alter the data before importing it into MongoDB. We used phpMyAdmin and MySQL to export the data from our Titanic datafile into a CSV-file.

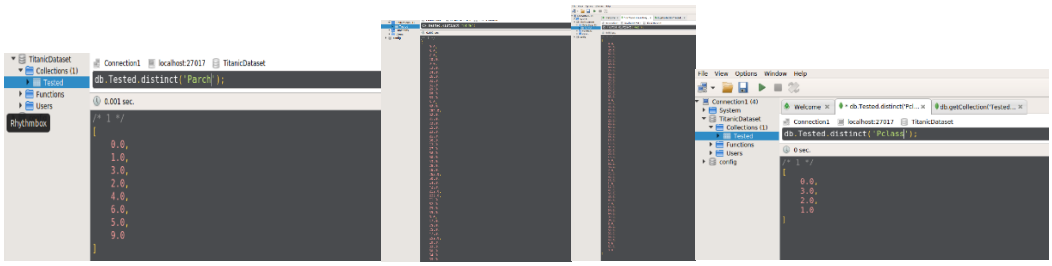
The titanic dataset file was transferred directly from MySQL into a MongoDB-friendly format in the third stage. The purpose of this phase is to change the dataset from MySQL into something MongoDB can use. The dataset will be inserted into a MongoDB instance as the last phase of this migration. We used the Ubuntu command line tool to import the Titanic dataset into MongoDB straight from a CSV-file.

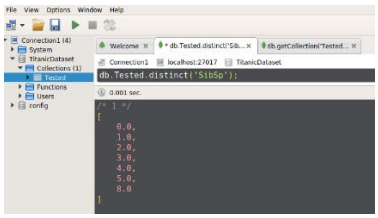
We were able to perform the same operations on NoSQL data that we performed in task 1. We used commands to find the lowest, maximum, average, and number of unique values per future, as shown below. The titanic dataset, which gives us information on every passenger, is seen in the first shot below. The additional photographs show that the execution of task 1 was completed successfully.

Results of Max ,MIN, Average and Unique



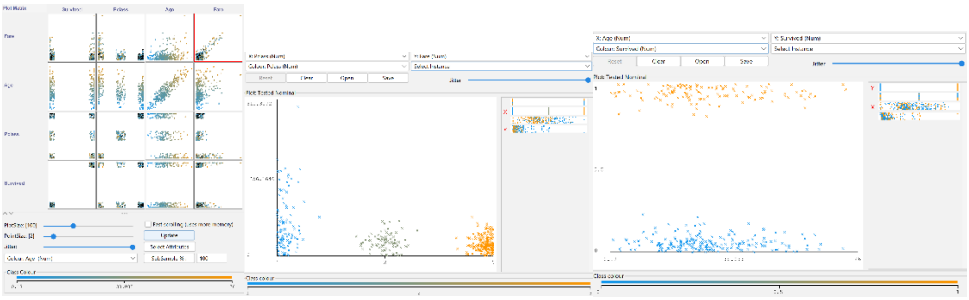




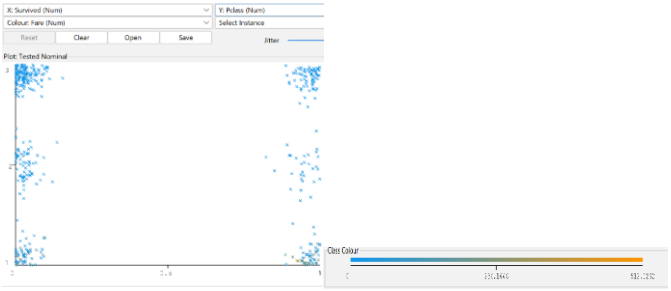


Task 3 –

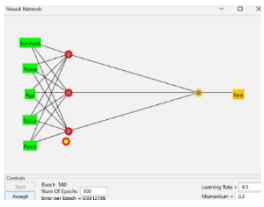
For our dataset we have selected the visible nominal Attributers, which narrowed our dataset from 10 to 4 columns. Deleting the unwanted and renaming our file. After doing so we have analysed our nominal dataset on Weka which gave us the results shown below. Weka helped us to analyse and see how the different attributes show different results when correlated.



Overall, the feature who has the largest influence is “Fare” for our classification problem, as it does show the relation between the other data. The reason for that being, we can see that people who paid most for their ticket survived the most and those who paid least did not. While it describes their age and so on. After training Neural Network and k-NN models we can see that “Fare” has most relation to our dataset, while the other attributes have relations with some and not all.



Results of running Neural Network and K-NN



Bibliography

Tambe, S., 2019. *What Are the Types of Cloud Models?.* [Online]
Available at: https://www.cuelogic.com/blog/3-types-of-cloud-computing-services
[Accessed 21 4 2022].

EDUCBA, 2020. *On Premise vs Cloud | Explore the Differences of On Premise vs Cloud.* [Online]
Available at: https://www.educba.com/on-premise-vs-cloud/
[Accessed 21 4 2022].

GeeksforGeeks, 2021. *Difference between Cloud Computing and Traditional Computing - GeeksforGeeks.* [Online]
Available at: https://www.geeksforgeeks.org/difference-between-cloud-computing-and-traditional-computing/
[Accessed 21 4 2022].

Vennam, S., 2020. *What is Cloud Computing?.* [Online]
Available at: https://www.ibm.com/cloud/learn/cloud-computing
[Accessed 21 4 2022].

Boisvert, M., Bigelow, S. J. & Chai, W., 2020. *Infrastructure as a Service (IaaS).* [Online]
Available at: https://www.techtarget.com/searchcloudcomputing/definition/Infrastructure-as-a-Service-IaaS
[Accessed 22 4 2022].

Tambe, S., 2019. *What Are the Types of Cloud Models?.* [Online]
Available at: https://www.cuelogic.com/blog/3-types-of-cloud-computing-services
[Accessed 22 4 2022].

Azure Microsoft, 2022. *Benefits of Cloud Migration—Why Cloud Computing? | Microsoft Azure.* [Online]
Available at: https://azure.microsoft.com/en-us/overview/cloud-migration-benefits-challenges/#overview
[Accessed 23 4 2022].

Liquid State, 2019. *The Advantages of PaaS: Leveraging a Platform to Add Value.* [Online]
Available at: https://liquid-state.com/advantages-paas-platform-as-a-service/
[Accessed 24 4 2022].

Wikipedia Contributors, 2022. *Platform as a service.* [Online]
Available at: https://en.wikipedia.org/wiki/Platform_as_a_service
[Accessed 24 4 2022].

Meegan, J., 2016. *Practical guide to PaaS: Benefits and characteristics.* [Online]
Available at: https://www.ibm.com/blogs/cloud-computing/2016/08/22/paas-benefits-characteristics/
[Accessed 24 4 2022].

IBM, 2020. *Top 5 Advantages of Software as a Service (SaaS).* [Online]
Available at: https://www.ibm.com/cloud/blog/top-5-advantages-of-software-as-a-service
[Accessed 24 4 2022].

McKenna Consultants, 2021. *The Relationship Between IoT, Big Data and the Cloud.* [Online]
Available at: https://www.mckennaconsultants.com/relationship-between-iot-big-data-and-the-cloud/
[Accessed 25 4 2022].

WebTribunal, 2022. *Cloud Adoption Statistics for 2022.* [Online]
Available at: https://webtribunal.net/blog/cloud-adoption-statistics/#gref
[Accessed 4 2022].

Citrix.com, 2021. *What Is Cloud Migration? How Does Cloud Migration Work? - Citrix Indonesia.* [Online]
Available at: https://www.citrix.com/en-id/solutions/vdi-and-daas/cloud-migration/what-is-cloud-migration.html?gclid=Cj0KCQjwgMqSBhDCARIsAIIVN1XrZTrEeCIiZV66hFZaz-mtZzpM7L45w0X4iGnhn4nPKyWI8gbXr2kaAgS_EALw_wcB&gclsrc=aw.ds
[Accessed 4 2022].

Vishnyov, D., 2021. *7 Essential Threats and Risks in Cloud Migration.* [Online]
Available at: https://itoutposts.com/blog/7-essential-threats-and-risks-in-cloud-migration/
[Accessed 4 2022].

Beacon Telecom, 2018. *Cloud Migration Strategies: The 6 R's.* [Online]
Available at: https://www.beacontelecom.com/cloud-migration-strategies-the-6-rs/
[Accessed 4 2022].

Jackson, C., 2022. *Cloud Provider Comparisons: AWS vs Azure vs GCP - Storage.* [Online]
Available at: https://acloudguru.com/blog/engineering/storage-showdown-aws-vs-azure-vs-gcp-cloud-comparison
[Accessed 4 2022].

A Cloud Guru, 2022. *Cloud Provider Comparisons: AWS vs Azure vs GCP – Compute.* [Online]
Available at: https://acloudguru.com/videos/cloud-provider-comparisons/cloud-provider-comparisons-aws-vs-azure-vs-gcp-compute?utm_campaign=11244863417&utm_source=google&utm_medium=cpc&utm_content=469352928666&utm_term=_&adgroupid=115625160932&gclid=Cj0KCQjwmPSSBhCNARIs
[Accessed 4 2022].

Jackson, C., 2022. *Cloud Provider Comparisons: AWS vs Azure vs GCP - Storage.* [Online]
Available at: https://acloudguru.com/blog/engineering/storage-showdown-aws-vs-azure-vs-gcp-cloud-comparison
[Accessed 4 2022].

Intel, 2014. *Cloud Architecture.* [Online]
Available at: https://www.intel.com/content/www/us/en/cloud-computing/cloud-architecture.html
[Accessed 4 2022].

Arcitura, 2022. *Platform-as-a-Service (PaaS) | Arcitura Patterns.* [Online]
Available at: https://patterns.arcitura.com/cloud-computing-patterns/basics/cloud-delivery-models/platform_as_a_service
[Accessed 4 2022].

Domb, M., 2019. *Smart Home Systems Based on Internet of Things.* [Online]
Available at: https://www.intechopen.com/chapters/65877
[Accessed 4 2022].

Turco, K., 2021. *4 Ways Cloud Computing Can Save Money | TechnologyAdvice.* [Online]
Available at: https://technologyadvice.com/blog/information-technology/4-ways-cloud-computing-can-save-money/
[Accessed 4 2022].

Globaldots, A., 2018. *13 Benefits of Cloud Computing for Your Business | GlobalDots.* [Online]
Available at: https://www.globaldots.com/resources/blog/cloud-computing-benefits-7-key-advantages-

for-your-business/
[Accessed 4 2022].

GeeksforGeeks, 2021. *7 Most Common Cloud Computing Challenges - GeeksforGeeks.* [Online]
Available at: https://www.geeksforgeeks.org/7-most-common-cloud-computing-challenges/
[Accessed 4 2022].

Patel, R., 2020. *10 Biggest Cloud Computing Challenges in 2022 for IT Service Providers.* [Online]
Available at: https://www.mindinventory.com/blog/cloud-computing-challenges/
[Accessed 4 2022].

SADOS, 2019. *The 5 Benefits of AWS Cloud Servers (And the 3 Drawbacks) | SADOS.* [Online]
Available at: https://sados.com/blog/aws-benefits-and-drawbacks/
[Accessed 4 2022].

Gaur, C., 2020. *Top 6 Big Data Challenges and Solutions to Overcome.* [Online]
Available at: https://www.xenonstack.com/insights/big-data-challenges
[Accessed 4 2022].

Gartner, 2012. *Definition of Big Data - Gartner Information Technology Glossary.* [Online]
Available at: https://www.gartner.com/en/information-technology/glossary/big-data
[Accessed 4 2022].

Simplilearn, 2019. *Big Data In 5 Minutes | What Is Big Data?| Introduction To Big Data |Big Data Explained |Simplilearn.* [Online]
Available at: https://www.youtube.com/watch?v=bAyrObl7TYE
[Accessed 4 2022].

Reliable Software Inc, 2019. *Using Big Data to Help Retailers Improve Their Business.* [Online]
Available at: https://www.youtube.com/watch?v=qwBS_IXw3JQ&ab_channel=ReliableSoftwareInc
[Accessed 4 2022].

Chan, M., 2018. *Cloud and Big Data: Why Cloud Computing is the Answer to Your Big Data Initiatives.* [Online]
Available at: https://www.thorntech.com/big-data-in-the-cloud/
[Accessed 4 2022].

Xenonstack, 2020. *Top 6 Big Data Challenges and Solutions to Overcome.* [Online]
Available at: https://www.xenonstack.com/insights/big-data-challenges
[Accessed 30 4 2022].

Wikipedia Contributors, 2021. *Bell–LaPadula model.* [Online]
Available at: https://en.wikipedia.org/wiki/Bell%E2%80%93LaPadula_model
[Accessed 1 5 2022].

Gaur, C., 2021. *Big Data Security Management: Tools and its Best Practices.* [Online]
Available at: https://www.xenonstack.com/blog/big-data-security
[Accessed 1 5 2022].

HG.org, 2022. *Security Breach - How Businesses May Be Liable - HG.org.* [Online]
Available at: https://www.hg.org/legal-articles/security-breach-how-businesses-may-be-liable-44358
[Accessed 1 5 2022].