Global Journal of Networks and Applications

Global Journal of Networks and Applications

Aims and Scope

Global Journal of Networks and Applications welcomes research contributions, surveys and notes in all areas relating to computer networks and applications thereof. The following list of sample-topics is by no means to be understood as restricting contributions to the topics mentioned:

- ➤ New design techniques, interesting or novel applications, components or standards.
- ➤ Interface issues including special consideration for handicapped persons.
- ➤ Computer graphics, 3-D modeling and virtual reality multi-and hypermedia including electronic publishing and digital libraries.
- ➤ Computer networks with tools such as WWW or Hyper wave.
- ➤ Emerging standards for internet presentation levels (such as XML) and Internet protocol level, new compression standards for still pictures, movies, audio and vector data, 3-D data and cartographic data.
- ➤ Work on metadata and its applications.
- ➤ Applications of networked and stand-alone multimedia systems to computer assisted presentations.
- ➤ Applications of an educational, transactional and co-operational nature.
- ➤ Gateways between databases and security, privacy and societal aspects of network and computer technology.

Global Journal of Networks and Applications

Global Journal of Networks and Applications

(Volume No. 11, Issue No. 2, May - August 2023)

Contents

Recent Trends and Developments in Computer Networks: A Literature Survey

[1] J. Ajay Kumar, [2] Girish P. Kulkarni, [3] Jyoti R Munavalli

[1,2]*UG student, Electronics & Communication Engg., BNM Institute of Technology, Bangalore*
[3]*Associate Professor, Electronics & Communication Engg., BNM Institute of Technology, Bangalore*
E-mail: [1]girishkulkarni1998@gmail.com, [2]kumarj855@gmail.com, [3]jyotirmunavalli@bnmit.in

A B S T R A C T

The field of Computer networks has grown tremendously in the recent past and still has scope to grow and create a lot of opportunities in this area. In this paper, we highlight the recent trends in computer networks and their applications. A literature review was carried out on different topics and synthesized. The survey on recent trends in computer networks depicts the effectiveness of these technologies in present and future in simplifying the problems of mankind.

Keywords- Cloud Networks, Learning, Automation, Artificial Intelligence.

I. INTRODUCTION

A computer network is a system in which multiple computers are connected to each other to share information and resources. The computer networks use open system interconnection (OSI) model. Protocol is the set of rules or algorithms which define the way how two entities can communicate across the network and there exist different protocols defined at each layer of OSI model [1,2].The protocols used are TCP/IP[1], UDP[2],HTTP[3],FTP[4] and other low level internet protocols such as ARP[5],ICMP[6]. Computer networks find application in information and resource sharing, speedy interpersonal communication, E – commerce and VoIP.

Nowadays networks are widely used everywhere. It's observed that a lot of innovations are happening in the field of computer networking and the main goal being the network security. The recent trends in the field of networking which are revolutionizing the digital world are Edge computing, Internet of Things (IoT), Cloud networking, Software Defined Network (SDN) Technology, Intent based learning, Artificial Intelligence, 4D Network, 5G, IPv6, automation and orchestration and many others. These technologies are being developed to have efficient and secured communication. In this paper, we perform literature survey on recent trends and developments in the area of computer networks.

---

[1]*Transmission Control Protocol/Internet Protocol*
[2]*User Datagram Protocol*
[3]*Hypertext Transfer Protocol*
[4]*File Transfer Protocol*
[5]*Address Resolution Protocol*
[6]*Internet Control Message Protocol*

II. MATERIALS AND METHOD

A literature survey was carried out in databases like Google scholar and web of science. The survey showed that the recent trends are in edge computing, cloud networking, software defined WAN and architecture, Cloud automation and orchestration, Intent based network, 5G and its applications, 4D network and Ipv6.

Edge Computing

Edge computing is a fragmented data processing that is specifically designed to handle the data generated by the Internet of Things [3]. Edge computing has mostly been to ingest process, store and send data to cloud systems. Only relevant data is separated from unwanted data and is sent to the network. The edge in edge computing refers to computing infrastructure that is close to the sources of data. For example: industrial machines (e.g. wind turbine, magnetic resonance (MR) scanner and undersea blowout preventers), industrial controllers such as SCADA systems, and time series databases aggregating data from a variety of equipment and sensors Edge computing systems are packing more compute, storage and analytic power to consume and act on the data [4]. They find application in industrial organizations and are more valuable. The ability to do on device processing and analytics is referred to as edge computing. Edge computing has new possibilities in IOT applications, Machine Learning for tasks such as face recognition, object detection, language processing[5, 6].It is estimated by 2020, the investment on edge infrastructure would be 18% of the total IOT infrastructure spent. The driving factors for edge computing are as follows:

- Faster processing
- Increase in cloud adoption
- Increase pressure on networks

The industrial machines collect huge amount of data that needs architecture with necessary goals: speed, remote contexts, saving bandwidth, storage, time, and costs by limiting the data that needs to be transmitted. The several key drivers making edge computing a more feasible reality today:

- Plunge in the cost of compute and sensors.
- More computing power executed in smaller footprint devices (such as gateway or sensor hub)
- Huge volume of data
- Machine learning and analytics

The edge computing will become critical in the following cases:

- Low connectivity in remote locations
- Bandwidth and associated high cost of transferring data to the cloud
- Low latency ,such as closed loop interaction between machine insights and actuation
- Access to temporal data for real time analytics
- Compliance ,regulation or cyber security constraints.[7]

Cloud Networking

Cloud computing is a rapidly growing technology and as per the latest Gartner report, it is projected that cloud services will grow at the rate of 17.3% in 2019. It is also estimated that by 2022, 90% of the organizations will be accessing cloud services. This is so because of its various advantages offered by these intelligent technologies. Cloud services are cheap; provide flexibility in accessing data anywhere anytime just with the help of good internet connection, ensures security of the data and has many more uses. Cloud networking is the method of sourcing or utilization of one or more network resources (shared networks) and services from the cloud. Cloud networking basically means sharing of network resources in a shared network whereas Cloud computing involves just sharing of computer resources. Examples of some network resources include virtual routers, firewalls, and bandwidth and network management software to name a few[8].

There are different types of cloud computing services and they are categorized into 3 types:

- **Infrastructure as a Service(IaaS)**: This service provides the infrastructure like Servers, Operating Systems, Virtual Machines, Networks, and Storage on rent basis. In this service organizations use their own platform and applications within a service provider's infrastructure. The user has to just pay on rental basis to the vendor for accessing the computing resources such as   servers, storage and networking. The largest cloud service providers Amazon Web Service and Microsoft Azure use this service.

- **Platform as a Service (PaaS):** This service is used in developing, testing and maintaining of software. PaaS is same as IaaS but also provides additional tools like DBMS and BI service. As the name suggests, here the service provider offers access to a cloud-based environment in which users can build and deliver applications. The provider supplies underlying infrastructure. The advantage of this service over IaaS is that it not only provides access to computer resources but also it comes with a suite of prebuilt tools to develop, customize and test their own applications. Apprenda, Red Hat Open Shift are the cloud providers which use this model.

- **Software as a Service (SaaS):** This service makes the users connect to the applications through the Internet on a subscription basis and accessing is done via the web or vendor APIs. In this service users do not have to manage, install or upgrade software, they can simply access the application through remote cloud network. Google Applications, Sales force use SaaS.

There are various cloud services providers throughout the world such as Kamatera, Phoenix NAP, Amazon Web Services, Microsoft Azure, Google Cloud Platform, Adobe to name a few top services. Recent studies suggest that AWS (Amazon Web Services) has emerged the largest cloud provider. It has become more popular because of the features like mobile friendly features, security, serverless cloud functions which is believed to improve the efficiency of any organization. Some of the cloud computing trends in 2019 are Hybrid & Multi-Cloud Solutions, Backup and Disaster Recovery, Serverless Computing, Data Containers, Artificial Intelligence Platforms, Cloud Security. Since public cloud offers very little options users required enhanced clouds and Hybrid cloud is one such cloud which is a mix of on-premises, private cloud and third-party, public cloud services with orchestration between the two platforms. In Serverless computing, the customer does not need to buy or rent services before and doesn't need to configure them. The Cloud is responsible for providing the platform, its configuration, and a wide range of helpful tools for designing applications, and working with data. This is a PaaS type of cloud service.

Artificial Intelligence (AI) is a fast growing trend in recent times. AI is implemented in those places which requires lot of automation and requires faster outputs. AI requires Machine Learning which in turn is developed from Artificial Neural Networks. AI can help to reduce costs in a variety of ways, such as preventing the duplication of effort, and taking over some expensive labor tasks, such as copying or extraction of data. The above theory suggests that Cloud computing has lot of uses and benefits and it really is a boon for the future technologies to be developed[9, 10].
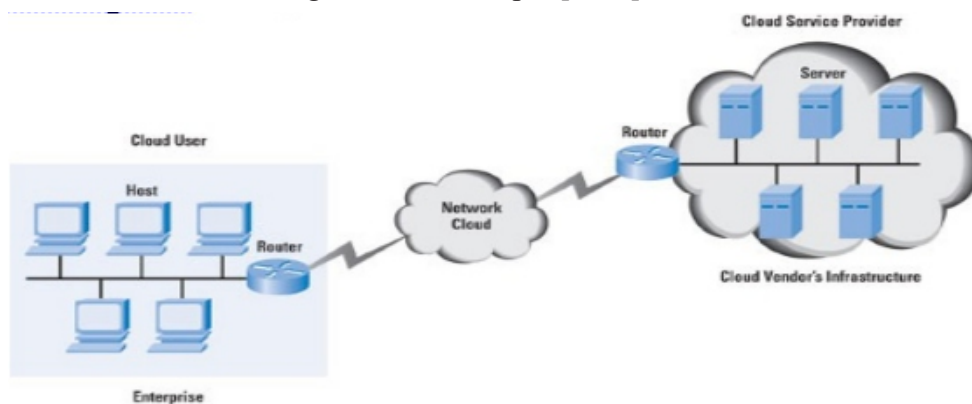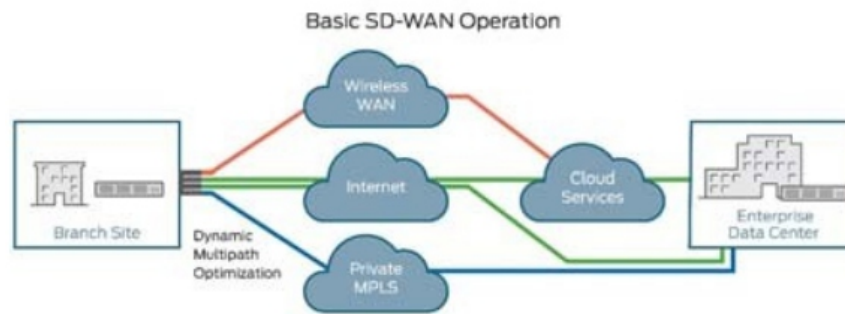

Fig. 1.Cloud Networking

Fig. 2. Software-defined WAN

Software Defined WAN (SD-WAN)

The global SD-WAN market was projected at $676.9 million in 2017 and is forecasted to witness the compound annual growth rate (CAGR) of 54.1% during 2018–2023. The demand for cost effective WAN management solutions, increasing adoption of cloud technologies, and the need for simplified network architecture as well as end-to-end network security and visibility are driving the market growth globally. This clearly explains the need for advance technologies like cloud computing, SD-WAN, SDN and many other technologies and SD-WAN is one prime example driving the technology market successfully.

CISCO defines 'SD-WAN as a software-defined approach to managing the wide-area network, or WAN', in a short and simple phrase. Actually there are many things involved in SD-WAN if we dig deeper into the concept [11, 12]. SD-WAN is a virtual WAN architecture that allows enterprises to leverage any combination of transport services – including Multiprotocol Label Switching (MPLS), LTE and broadband internet services – to securely connect users to applications. A wide area network (WAN) is a network that exists over a large-scale geographical area. A WAN connects different smaller networks, including local area networks (LANs) and metro area networks (MANs). TCP/IP is the protocol used for a WAN in combination with devices such as routers, switches, firewalls and modems.

The reason why SD-WAN came into picture and took over conventional WAN is that   traditional WANs based on conventional routers are not cloud- friendly and unlike the traditional router- centric WAN architecture, the SD-WAN model is designed to fully support applications hosted in on- premise data centers, public or private clouds and SaaS solutions such as Salesfore.com, Workday, Office365 and Dropbox, while delivering the highest levels of application performance which enhances business operations. Another problem with traditional WAN was backhauling of traffic importantly. Backhauling simply adds latency or delay that impairs application performance resulting in poor user experience and loss in productivity where users expect speedy computations especially in business environments. All these were overcome by using SD-WAN[13].

Software defined Networking (SDN)

It has been forecasted that the global Software Defined Networking Market size to grow from 8.8 billion USD in 2018 to 28.9 billion USD by 2023, at a Compound Annual Growth Rate (CAGR) of 26.8% during the forecast period. SDN is emerging as one of the most lucrative and promising networking technologies, which has the competence to enable network virtualization and innovation, addressing most of the persistent networking problems. Software defined Networking (SDN) is the physical separation of the network control plane from the forwarding plane, and where a control plane controls several devices.

SDN is known for its innovative approach of separating the data plane from the control plane, a clear improvisation to the traditionally distributed control planes. This is executed by splitting the hardware of Ethernet switches (data plane) from the logical link that controls the flow of packets (control plane).The data plane also known as the user plane, forwarding plane, carrier plane or bearer plane is the part of a network that carries user traffic. The data plane, the control plane and the management plane are the three basic components of telecommunication architecture [14]. The control plane and management plane serve the data plane, which bears the traffic that the network exists to carry. Data plane refers to all the functions and processes that forward packets/frames from one interface to another. Control plane refers to all the functions and processes that determine which path to use. Routing protocols, spanning tree, ldp, etc are examples for the same. Management plane is all the functions you use to control and monitor devices.

SDN-Architecture:

A typical representation of SDN architecture comprises three layers: the application layer, the control layer and the infrastructure layer (Fig.3).

- The application layer contains the typical network applications or functions organizations use, which can include intrusion detection systems, load balancing or firewalls. Where a traditional network would use a specialized appliance, such as a firewall or load balancer, a software-defined network replaces the appliance with an application that uses the controller to manage data plane behavior.
- The control layer represents the centralized SDN controller software that acts as the brain of the software-defined network. This controller resides on a server and manages policies and the flow of traffic throughout the network.
- The infrastructure layer is made up of the physical switches in the network.

These three layers communicate using respective northbound and southbound application programming interfaces (APIs).

Software-defined networking uses an operation mode that is sometimes called adaptive or dynamic, in which a switch issues a route request to a controller for a packet that does not have a specific route. This process is separate from adaptive routing, which issues route requests through routers and algorithms based on the network topology, not through a controller[15, 16].
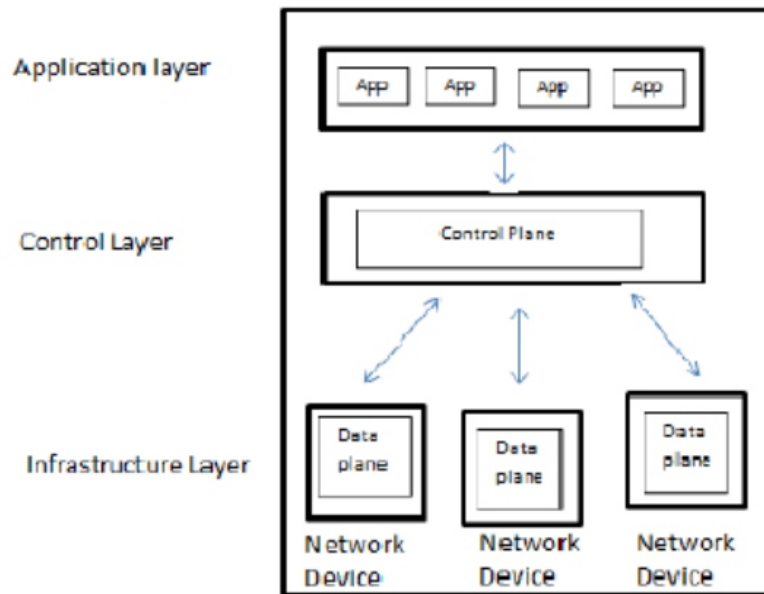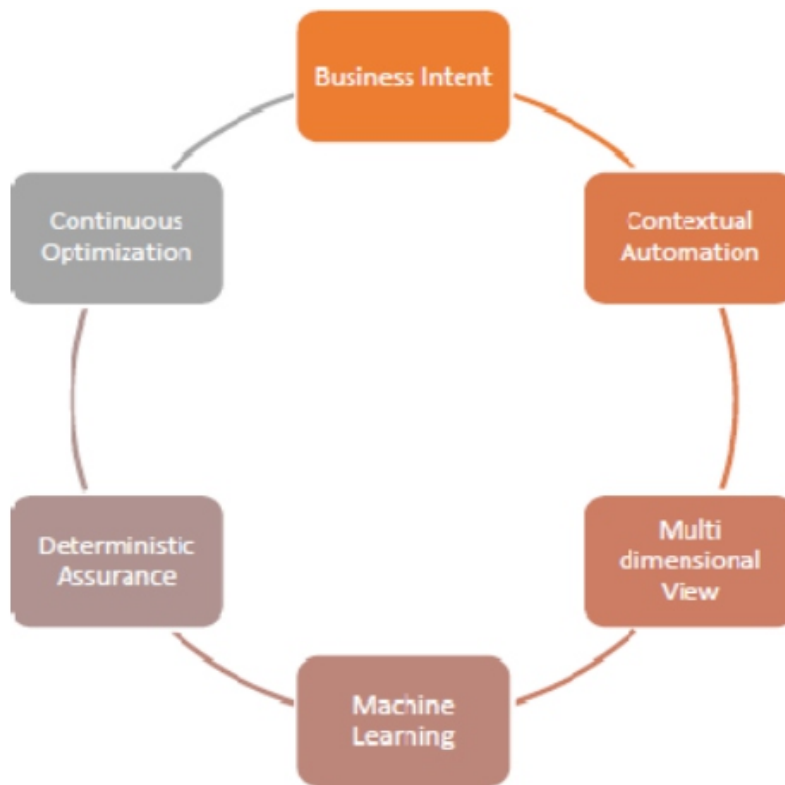

Fig.3. Layers in SDN Architecture

Fig.4. Intent based Networking

Cloud Automation and Orchestration

The global cloud orchestration market is estimated to be valued at nearly 6,970 mn USD in 2017 and is projected to cross 20 bn USD by 2025 end. Sales revenue is expected to register a CAGR of 14.6% during the forecast period (2017–2025). This indicates that in the coming years, it is expected that small and medium enterprises are going to be top adopters of cloud orchestration platforms and services. Cloud automation refers to deployment and management actions that IT teams can carry out without human interaction to say the least. Cloud orchestration directly relates to cloud automation, as it arranges, sequences and coordinates automated tasks. Cloud orchestration is the use of programming technology to manage the interconnections and interactions among workloads on public and private cloud infrastructure.Cloud orchestration is typically used to provision, deploy or start servers; acquire and assign storage capacity; manage networking; create VMs; and gain access to specific software on cloud services. This is accomplished through three main, closely related attributes of cloud orchestration: service, workload and resource orchestration. An orchestration platform can integrate permission checks for security and compliance.

Cloud orchestration gained importance since 2006 when IaaS andPaaS,multi-cloud deployments went on to become massive in scale and hence Management of a cloud infrastructure is becoming a tremendous challenge for in-house IT staff. Here the cloud orchestration plays very important role where its tools are used to simplify deployment/management and provide added visibility into cloud resources and processes. Orchestration software combines all the possible cloud resources, and then provides a mechanism to link and automate the provisioning for various services organizations require. Orchestration brings high availability, scaling, failure recovery, dependency management, and numerous other tasks and attributes into a single process that can tremendously reduce staff effort. Orchestration also provides visibility into resources and processes that simple cloud automation lacks. Some of the cloud orchestration vendors and their products are Amazon Web Services Cloud Formation, Cisco Cloud Center.

Automation is a subset of orchestration, which means that orchestration provides coordination among and across many automated activities. Scripts are being used to automate deployment. Scripts are a flexible way to write the commands as a file that can be quickly executed. Two popular open source automation tools are Chef and Puppet. The Chef model is procedural because it explicitly describes how to deploy and connect cloud application components. Chef creates "recipes" and "cookbooks"that are built in a programmer-friendly way. Each deployment step can be described independently, and bringing those independent recipes together creates a repeatable application deployment process. Since every operations step can be described in a recipe, there is nothing that can be deployed manually that

Chef can't automate. If Chef is programmer-friendly, Puppet's origins lie with operations people, who are more familiar with describing what they want than the detailed steps in getting there. Like a high-level programming language, Puppet's models can simplify deployment descriptions. Furthermore, because an end-state approach can be used to describe any step in application lifecycle management (ALM). Overall this is how cloud automation is implemented. Cloud automation and orchestration work together, and one is dependent upon the other. It doesn't make sense to create an orchestration process that calls no automated tasks, nor would it make sense to create automated tasks that could not be a part of a cloud orchestration process [17,18].

Intent Based Networking:

Intent Based Networking is the most developing area of network technology which uses machine learning and artificial intelligence to automate network management (Fig. 4). IBN looks forward in developing self managing networks .Most of the current networks are managed manually. Network administrators create scripts which has every detail required to accomplish a given task and interact through a command line interface and program the switches and scores of other devices supporting enterprise networks. Intent based networking is giving commands to the network on what you want (intent) rather than what to do, to accomplish the expected goal. The intent is interpreted to all the devices in the network automatically without requiring engineers altering each device in the network. The IBN needs fewer inputs which are through a user friendly graphical user interface (GUI) [19].

Intent based networking system is an adaptive system that learns from the previous situations using artificial intelligence and would respond to the question of what should be done in a given situation. The response to a given situation could be an automatic reaction to some system event. Given an event or a desired event the system could work out how to respond appropriately without human interference. Intent Based Networking finds application in data centers, public clouds and internet of things. Since AI is still developing stage, the IBN is also in early stages of development. IBN will be functioning in most of the networks in the near future as the world is moving towards Artificial Intelligence.

5G and its applications

The needs for the development of fifth generation of network technology are as follows:
- Increase in demand wireless broadband services
- Faster Wireless Networks
- High capacity networks
- Internet of Things (Smart Devices)
- Low Latency Networks
- Reliable networks

The field of technology required to develop 5G Networks are

- Radio Access Technology
- Antennas
- Network Slicing
- Air interface Design (spectrum utilization)

Sl.No.	Parameter	Value
1	Latency in the air link	<1ms
2	Latency end-to-end (device to core)	<10ms
3	Connection density	100X compared with LTE
4	Area capacity density	1Tbits/s/km^2
5	System spectral efficiency	10bit/s/Hz/cell
6	Peak throughput (downlink) per connection	10Gbit/s
7	Energy efficiency	>90% improvement over LTE

Table 1: Parameter values

5G is advanced in terms of:

- High increased peak bit rate
- Larger data volume per unit area (i.e. high system spectral efficiency)
- High capacity to allow more devices connectivity concurrently and instantaneously
- Lower battery consumption
- Better connectivity irrespective of the geographic region, in which you are
- Larger number of supporting devices
- Lower cost of infrastructural development
- Higher reliability of the communications

Applications of 5G include in the area of IoT, Smart farming, Logistics and shipping, High speed network, Fleet management, Smart Cities, Autonomous Driving, Security and Surveillance. [20]. Fig. 5 shows different development areas of 5G technology and LTE-A.
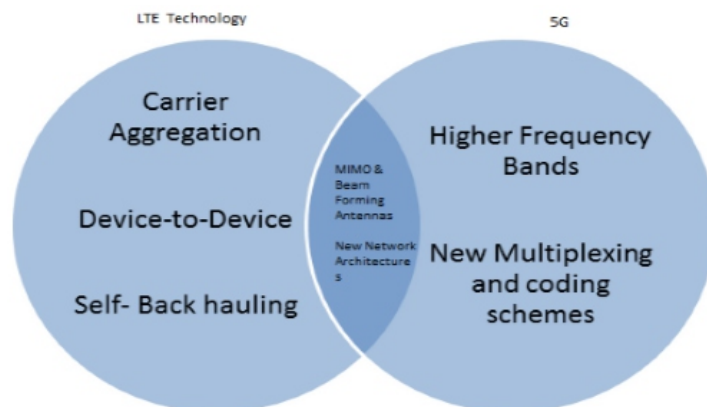

Fig.5 Development areas of LTE-A and 5G

technology
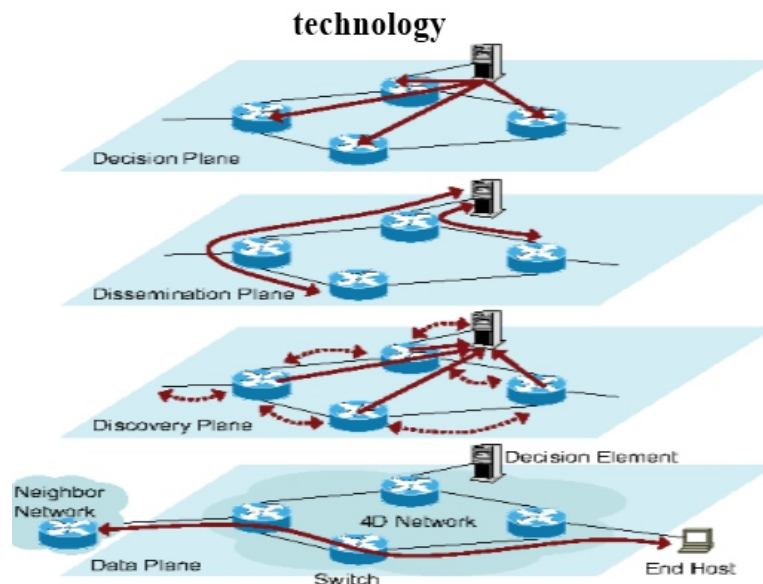
Fig. 6. Planes of 4D network

4D Network

Tesseract is an experimental system that enables direct control of computer network under a single administrative domain. Tesseract is based on 4D Architecture, where the network control plane is decomposed into decision, dissemination, discovery and data planes (Fig. 6). It provides two important services to enable direct control: the dissemination service that carries control information from the network decision element to the nodes of the network and node configuration service provides an interface for the decision element to command the nodes in the network to perform desired operations. Tesseract reduces human error by minimizing the need for manual node configuration. It can handle up to thousand nodes in intra-domain routing control.Tesseract flexibility is shown by its application in joint packet forwarding and policy- based filtering for IP networks and in link-cost driven Ethernet packet forwarding [21, 22].

The planes of 4D Architecture are as follows:

Data Plane: This plane operates in network switches and provides services such as IPv4, IPv6,Ethernet packet forwarding. The actions of data plane are based on states in the switches, and this state is controlled by decision plane. Example state in switches includes the forwarding table or forwarding information base, queue-management parameters, tunnels and network address translation mappings, flow scheduling weights, packet filters etc.

Discovery Plane: Each switch is responsible for discovering its hardware capabilities and its physical connectivity to neighboring switches. The information obtained is then reported to decision element in

decision plane through logical connections maintained by dissemination plane. End hosts do not participate in discovery plane for backward compatibility.

Dissemination Plane: This plane maintains the logical channels that carry control information between network switches and decision plane. The dissemination paths are separately maintained from data paths so that they can operate without requiring reconfiguration or successful establishment of paths in data plane. The today's network's control and management information is carried over data paths, which has to be accepted by routing protocols, which in turn creates circular dependency.

Decision Plane: It has a logically centralized decision element that makes all decisions driving network control such as access control, load balancing, and security. The information obtained by the discovery plane is used by the decision element to make decisions and these decisions are sent as commands to switches via dissemination plane. The node configuration service interface helps the decision element to command the switches.

Internet Protocol version 6 (Ipv6)

IPv6 is the most recent version of the Internet Protocol, the protocols help in identification and location of the system over the network and route traffic across the network. The Internet Engineering Task Force (IETF) developed IPv6 to deal with the problems of IPv4 address exhaustion.

Ipv6 will replace the current IPv4 devices for identification and location definition. Due to the rapid growth of internet in the recent decade, the addresses required connecting the devices increased and the IPv4 exhausted which led to the development of IPv6. IPv6 has a 128 bit address, theoretically allowing 2128 or approximately 3.4 X 1038 addresses. Some addresses are reserved for special use and the total number of available IPv6 addresses is more than 7.9 X 1028 times as many as IPv4, which uses 32 bit addresses and provides approximately 4.3 billion addresses. However, IPv6 allows communication between IPv4 and IPv6 hosts.(Fig. 7).
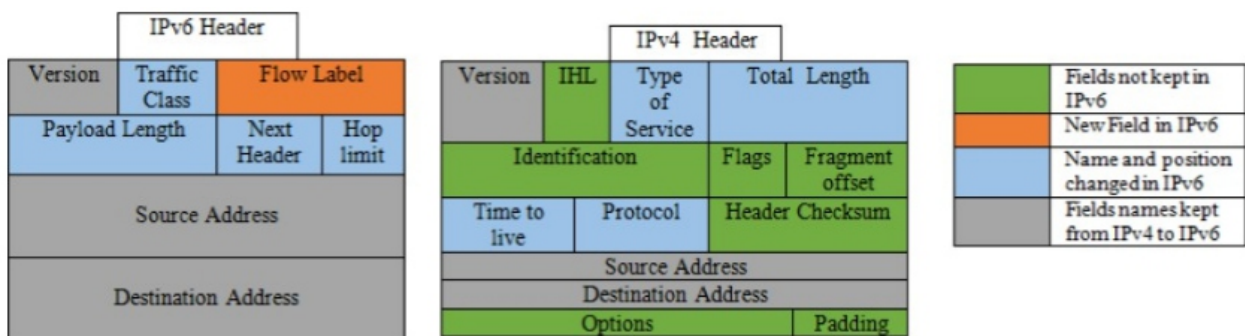

Fig.7 Comparison of IPv6 and IPv4

IPv6 has some technical benefits other than larger address space. It allows hierarchical address allocation methods that facilitate route aggregation across the internet and limit the expansion of routing table. IPv6 includes device mobility, security and configuration aspects.IPv6 is an Internet Layer Protocol for packet switched internetworking and provides end-to-end datagram transmission across multiple IP networks. It simplifies the aspects of network renumbering, address configuration, router announcements when changing network connectivity providers. In routers, IPv6 simplifies the processing of packets by placing the responsibility for packet fragmentation into the end points. Subnet size of IPv6 is standardized by fixing the host identifier portion of an address to 64 bits. Network security was the most important design requirement of IPv6 architecture [23, 24].

III. CONCLUSION

The discussed technologies have increased the efficiency and security of networks. The world is striving to make things easier for all. The applications of these technologies are wide spread and being used in every nook and corner of the country. The trending technologies are IoT, AI, Machine Learning and Cloud Computing. Additionally, big data and its analytics have opened new opportunities in computer networking. It is estimated that the above technologies will conquer the entire digital market and the world is becoming smart by using these innovative technologies.

REFERENCES
[1].www.geeksforgeeks.org
[2]. www.tutorialspoint.com
[3].networkworld.com
[4]. ge.com
[5]. Weisong Shi, JieCao, Quan Zhang, Youhuizi Li, and LanyuXu,'Edge Computing: Vision and Challenges', IEEE Internet of Things Journal, Vol.3.No.5, 2016.
[6]. Yuan Ai, Mugen Peng, Kecheng Zhang, 'Edge computing technologies for Internet of Things: a primer, Digital Communications and Networks', Vol. 4, No.2,2018,
[7]. Sonia Shahzadi, Muddesar Iqbal, Tasos Dagiuklas and Zia Ul Qayyum, 'Multi-access edge computing: open issues, challenges and future perspectives', Journal of Cloud Computing Advances, Systems and Applications, Vol 6, No.30, 2017.
[8]. Santosh Kumar and R. H. Goudar,' Cloud Computing – Research Issues, Challenges, Architecture, Platforms and Applications: A Survey', International Journal of Future Computer and Communication, Vol. 1, No. 4, 2012.
[9]. Prasad, M.Rajendra&Ramavathu, Lakshman & V, Dr. Bapuji. 'Cloud Computing: Research Issues and Implications', International Journal of Cloud Computing and Services Science. No.2,2013.
[10]. Dar, Ab & Ravindran, Daks, 'A Comprehensive Study on Cloud Computing', 2018.
[11]. William Stallings, 'Foundations of modern networking: SDN, NFV, QoE, IoT, and Cloud' 1st Edition, Pearson Education, Inc, 2016
[12]. Harpreet Singh, 'Implementing Cisco Networking Solutions',Packt Publishing Ltd., 2017.
[13]. Shailendra Mishra and Mohammed Abdul Rahman Al Shehri, 'Software Defined Networking: Research Issues, Challenges and Opportunities', Indian Journal of Science and Technology, Vol.10, No. 29, 2017.
[14]. Raphael Horvath, Dietmar Nedbal and Mark Stieninger, 'A Literature Review on Challenges and Effects of Software Defined Networking', Procedia Computer ScienceVol. 64, 2015.
[15]. William Stallings, 'Foundations of Modern Networking: SDN, NFV, QoE, IoT, and Cloud', 1st Edition, 2015.
[16]. Jayachandran, Prassanna, R Pawar, Anjali, Venkataraman and Neelanarayanan, 'A review of existing cloud automation tools', Asian Journal of Pharmaceutical and Clinical Research., 2017.

[17]. Omid Givehchi, Henning Trsek and Juergen Jasperneite, 'Cloud computing for industrial automation systems — A comprehensive overview', IEEE 18th Conference on Emerging Technologies & Factory Automation, 2013.

[18]. Dirk Schulz, 'Intent-based automation networks: Toward a common reference model for the self-orchestration of industrial intranets', 42nd Annual Conference of the IEEE Industrial Electronics Society IECON 2016.

[19]. Swati Yadav and Sugandha Singh, 'Review Paper on Development of Mobile Wireless Technologies (1G to 5G)', International Journal of Computer Science and Mobile Computing, Vol.7, No.5, 2018.

[20]. Meenal G. Kachhavay and Ajay P.Thakare, '5G Technology-Evolution and Revolution', , International Journal of Computer Science and Mobile Computing, Vol.3, No.3, 2014.

[21]. Yan, Hong &Maltz, David & S. Eugene Ng, T &Gogineni, Hemant & Zhang, Hui & Cai, Zheng, 'Tesseract: A 4D Network Control Plane', 4th Symposium on Networked Systems Design and Implementation, 2007.

[22]. Greenberg, Albert &Hjalmtysson, Gisli&Maltz, David & Myers, Andy & Rexford, Jennifer &Xie, Geoffrey & Yan, Hong & Zhan, Jibin& Zhang, Hui, 'A clean slate 4D approach to network control and management', ACM SIGCOMM Computer Communication Review. 35, 2005.

[23].Paul, Henry &Bakon, Kinn, 'A Study on IPv4 and IPv6: The Importance of Their Co-Existence', International Journal of Information Systems and Engineering, 4, 2016.

[24]. Sailan, Mohd.Khairil& Hassan, Rosilah& Patel, Ahmed, 'A comparative review of IPv4 and IPv6 for research test bed', Electrical Engineering and Informatics, Vol.2, 2009.

Application of an Artificial Neural Network to Predict Employment: A Case Study in A Bachelor of Sciences (Information Technology)

[1] Thirachit Saenphon, [2] Waralak V. Siricharoen

Faculty of Information and Communication Technology, Silpakorn University
E-mail: [1]thirachits@gmail.com, [2]lak_waralak@yahoo.com

A B S T R A C T

This research is an application of data mining techniques to predict a job that meets the academic qualification in bachelor's degree of information technology, Faculty of Information and Communication Technology, Silpakorn University. The results of the study in English, Core Program and the features needed to complete the course. There are 165 graduated students who completed the 2011 updated curriculum. This research uses two methods of data mining, which are divided into 2 phases: Decision Tree and Neural Network. The features needed are the knowledge of global economy, IT management, strategically, solution management, staff management and quality management, systems development process, programming and testing. Program that affects the job is the E-business, Management information systems, Human resource management, Senior project, Marketing, English language, Human computer interaction design and Object-oriented design.

Index Terms - Bachelor, Data Mining, Neural Network, Prediction

I. INTRODUCTION

The purpose of education at the undergraduate level is to prepare students for work or to study for a master's degree with personality preparation. Features are ready for work such as responsibility, interpersonal interaction, team work etc. [1] Those undergraduate level students have developed knowledge and the ability to apply the theory to practice. When students have chosen to study in the field/ faculty they preferred, some students may have problems in education such as low performances throughout the course. The cumulative grade point average (GPA) is an important factor for applying for jobs in accordance with the degree of information technology which has many job positions. These positions need the person who has the basic skills that match the job characteristics, such as programmers, it needs to have analytical skills to solve problems, programming skills in the language that the market wants.

If students are studying in the field of information technology get to know their tendency to prepare academic skills, personality skills to apply for jobs and get work that is related to the field of information technology. Changing learning behavior and preparing the necessary skills in a timely manner during their study. To make their grades high enough for the criteria that help them get a job that matches the graduation in the organization.

Therefore, the authors are interested in developing a model for predicting the tendency to get a job that meets the qualification of information technology using artificial neural networks. The results of the research can be used to develop a trend forecasting system to get the right work positions that meets the educational qualification. Also, the results can be used for students studying at the undergraduate level to continue to benefit the preparation of the information technology job market.

In this paper is the short overview of the ongoing research work in literature review and data preparation. The scope of study and research contribution also is mentioned in the paper. The paper is concluded in with a discussion and conclusion.

II. DATA MINING TECHNIQUES REVISITING

Important techniques of data mining are Data classification and Prediction.

A. Data Classification and Forecasting

1. Data Classification

It is the technique to separate information into categories according to the characteristics of each target group or class. The result will be a model for classifying, to be used to separate samples of unknown classes. The modeling process is divided into 3 phases:

Model Construction is a modeling process by learning from the information that has already been assigned to the class (Training data). The training data in the model can be displayed in the form of a decision tree, where each node displays the attribute in each branch shows the results of the test. And the Leaf node shows the result as the specified class. Another way to create a model is the Neural Network by mimicking the learning of cells in the brain. It is a method that allows computers to learn from model examples and train the system to know and solve problems in the test data.

Model Evaluation is the process of evaluating accuracy using test data. The actual class of data used in this test will be compared with the class that is calculated from the model. This is to test the accuracy.

Model Usage is a model for use with unseen data. The class will be assigned to the new data or predict the class value.

2. Data Forecasting

It is a prediction or prediction of the desired value from existing data. It is often used with predictions of continuous value characteristics, such as sales forecasts for the next month from existing data. Predict the disease from the patient's condition, etc.

B. Finding the relationships (Association)

Finding patterns that occur as the frequently links that occurred (Association) or correlation (Correlation) from various units of data in the database. This is the way to analyze or predict various phenomena such as product placement, in order to stimulate sales Cross-product marketing (Cross-marketing) planning, promotion, sales support etc.

C. Clustering

The grouping of data is divided into sections, each section is appropriately related. Without prior grouping (Unsupervised Learning) which is different from the classification of information. It is classification which was grouped by using a distance to indicate similarities and the difference of each row of data. The similar data rows have a small distance between rows of data. While different, data rows, there will be a lot of distance between the data in rows.

Examples of grouping, organizations want to know the similarities, which they have in their customers. So that the organization can understand the characteristics of the target customers and for the organization to develop products or sell products to meet customers' needs. The organization will separate the group of customers into groups by using the technique of clustering.

For this research, the authors have chosen to use predictive techniques. By using the Decision Tree method to find features and courses that influence getting a job that matches the educational qualification in information technology. And choosing the method of neural network to create a model for predicting the tendency that students can apply for jobs with the qualification of information technology.

III. LITERATURE REVIEW

Data mining techniques commonly used in educational data analysis such as prediction model for the suitability of new students in the field of information technology with relational classification rules with Class Association Rules (CARs). It is applying data from Information Technology Students, Chandrakasem Rajabhat University. There are 18 factors used in the research as follows: gender, age, age when studying, residence, and cumulative grade point average, funding status, aptitude, total Number of siblings who are studying, father and mother information, status, income, occupation [2].

Using data mining techniques to predict student outcomes, this research aims to develop a data warehouse and create a model for predicting student results by using secondary school students consists of 16 personal data and academic performance data as follows: gender, study plan, occupation

of father, average annual income of father mother's occupation average income per year of mothers status of parents, number of siblings, average academic results, mathematics average science results average results of Thai language average results of social studies average English results. The average grade level results of grade 1, average grade level results, grade 2, average grade level results, grade 3, and the output data or results that need to predict 1 feature is the average grade level at grade level 4 [3].

Neural networks were used by many researchers for predictions of students' results[4]. Cooper presents a neural network-based decision support system that identifies students who are "at-risk" of not retaining to their second year of study. The system correctly predicted retention for approximately 70% of the students.

Pannipa Boo-ek and Suradech Boonlue [5] have adopted machine learning algorithms that include vector machine support (SVM) techniques, decision tree techniques and outline techniques. Reverse neural network (BP-ANN) applied to create a model for the Oka prediction of students' graduation Then compare the efficiency, accuracy in forecasting. To find a forecast model that provides the highest accuracy that is suitable for use in forecasting the graduation opportunities of computer science students, Department of information technology North Bangkok University During the academic year 2004-2008, a total of 138 records, including personal history data and academic performance data of students Which is divided into 18 inputs, namely gender, income, father, income, mother, father's status, maternal status, average grade, original school, blood group, private school / government school Residential area Grant status Study group Number of years GPA 1 Semester Grade Semester 2 Summer GPA Cumulative grade point average, 1st semester, cumulative grade point average, 2nd semester and summer cumulative grade point average And 1 release information, which is the student's graduation status (Graduated according to the specified time / not completing the study according to the specified period). The data used in the research is the data of the first-year students only. The results of the test showed that the forecasting model that uses SVM-PK techniques to provide the highest predictive accuracy at a percentage 89.13 with an accuracy of 0.878 the memory value is 0.891 and the overall performance value is 0.860. In addition, the minimum RMSE value is 0.3297. With the forecasting model supporting vector machines that use kernel functions like SVM-NP SVM-RBF and SVM-PUK, providing the same forecast accuracy of 87.68 percent, C4.5 andBP-ANN forecasting models. The accuracy of 83.33 percent and 84.36 percent, respectively, can be concluded that the forecast model SVM-PK is the most suitable model for predicting student graduation opportunities.

Bogdan Oancea [6] studied important problems in higher education, namely Student results after attending university. Many students leave the University for Several Reasons, including basic

education and low education. The financial ability predicting student outcomes is a major problem for university management that is needed. To avoid premature out-of-school phenomena, Bogdan and the faculty use artificial neural networks to predict student outcomes by measuring the average score in the first year of study. By sampling 1000 students from the University of Bucharest from the latest version of the three graduates, 800 models for training in models and 200 people to test models' Neural network is Multi-layer perceptron (MLP) and trained using reverse diffusion algorithms (Back propagation) Import data is details of students while enrolling at the university, including information about age GPA with high school graduation. The gap between high school graduation and enrollment in higher education after the training of artificial neural network models the results of the MSE forecast of approximately 1.7% of the university's executives can be used to predict the results of the students in order to carry out early prevention to avoid premature exit phenomena. Neural Networks applied in data clustering problem. In [7] Moucary et al. study Neural Networks and Data Clustering based method to predict students' GPA according to their foreign language performance for those students who study in a foreign language. In the second step, the students will be grouped in the group that is already set up to give additional suggestions. This study has the highest predictive error of less than 10% for GPA.

IV. RESEARCH BACKGROUND

A. Undergraduate standards computer major in Thailand

Business Information Technology Program, Faculty of Information and Communication Technology, Silpakorn University [8]. This program is one of the 3 majors in the Faculty of Information and Communication Technology. It takes 4 years to complete each major. The course identifies a career that can be assembled after graduation. For example, Information Technology, Information system analyst, and designer, programmer, Website designer and developer Information Project Manager, Database administrator, Information Technology Manager Professionals in business establishments or business organizations. Because this program is under the identity of Silpakorn University, which is a university with a reputation for art. Therefore, there are program in art for students to study until having the ability to pursue a career in website design, computer graphics etc.

B. IT Skill Standards (ITSS)

In accordance with professional information technology standards, ITSS, which is an international cooperation led by Japan, Thailand, Philippines, Vietnam and many other countries in Asia. To create professional standards for information technology to be accepted internationally has categorized 11 categories of work, 35 specialization fields [9]. Fig. 1 is the framework of the profession defined in the

market according to the type of work. Which each position in each organization usually, the cumulative grade point average is determined. In recruitment including when the selection is in the consideration process there are also considerations of course level results. Skills that are essential to the profession that graduates have already learned. Therefore, for students who are studying to realize the importance of preparing their own future. The authors therefore used the information of graduates who graduated during the year 2013-2015, which received work that was directly related to Kutawut. To analyze to create a forecasting model for the tendency of getting a job that meets the educational qualification Information Technology (Office of Information Technology Promotion, 2016).

The program that the authors analyze the data is the 2011 course which is a graduate student that has graduated. The total number of credits in the program is 149 credits. The course structure consists of the following [8]:

General Education Courses	30 credits
Required courses	9 credits
Language group	
Elective subjects not less than	12 credits
(Each group of not less than 3 credits): Humanities, Social Sciences, Mathematics and Science	
Elective courses set by the faculty of	9 credits
Specific courses	113 credits
Core Courses	36 credits
Major subjects required	62 credits
Majors choose	15 credits
Free Elective Courses	6 credits

From Fig 1, the authors summarize the details of job positions to include only the main positions. To respondents only because in this sub-position according to ITSS, there are some positions that do not have a graduate work. Therefore, this may result in poor analysis and some sub-positions have names called in different workplaces, which may cause the questionnaire to not respond to reality. The authors then summarize the position according to Fig 1 to remain at the position of 11 main positions only.

1. Marketing (Marketing)
2. Sales (Sales)
3. Consultant
4. Architectural work in IT (IT Architect)
5. Project Management
6. Specialized IT (IT Specialist) jobs
7. Application Specialist
8. Software Development
9. Customer Service (Customer Service)
10. IT Service Management
11. Education

Fig. 1. The framework of the profession defined in the market according to ITSS (adapted from IT skill standard (ITSS))

C. Scope of study

1. Study from the sample group who are graduates in the Bachelor of Science program. Business Information Technology from the Faculty of Information and Communication Technology, Silpakorn University from the academic year 2013-2015 and get work done in 165 organizations operating in Thailand.

2. Study variables- Independent variables include personal information on the grades of graduates, including grades, grades in the subject, special subjects in elective subjects, foreign languages including scores from making a feature test that corresponds to the graduate characteristics of the information technology field. Variables based on job positions in specialized fields of information technology

This research uses the sample data used to test the data of graduates of the Bachelor of Science Program. Business Information Technology Branch Faculty of Information and Communication Technology, Silpakorn University (completed 2011) since the academic year ending 2014-2016, the number of graduates 166 people, 165 respondents. By students were entering the program in year 2014, ending the academic year 2017, with a total of 80 graduates (not collected since finishing after the experiment). The students were entering the program in year 2013, ending the academic year 2016, the number of graduates 84 people and students entering the year 2012, ending the academic year 2015, the number of graduates 82 people (including 2015-2016 number of graduates, 166 people), the authors collected 165 questionnaires for those who entered the year 2011, graduating the academic year 2014. The number of graduates was 83 people. Gets cooperation in answering queries considering the group of respondents. The purpose of the research by collecting personal information, information on the grades of graduates, including grades, grades in the subject, special subjects in elective subjects, foreign languages including scores from the feature tests that correspond to the graduates of information technology.

D. Research contribution

1. Educational institutions have information technology program can apply research results as guidelines for curriculum development. To develop student characteristics to meet the requirements of information technology positions.

2. Students are aware of the features specialized knowledge, which is suitable for the position of information technology that they want. In order to be prepared with the knowledge gained from the course of the institution that he studied and includes features Personality suitable for the job.

E. Knowledge, skills and competencies of information technology personnel of Thailand

For information technology personnel, the necessary knowledge, skills and competencies can be classified into 4 parts, which are related to planning. The part about creating or developing Parts related to use and related to general management (Study project to establish the standard of position determination to support government personnel and government information, 2013) with the following details:

Part 1: With respect to planning will have knowledge, skills and competencies as follows: Knowledge of ICT system design to match the organization's mission Knowledge of ICT management, both strategic Problem Solving, Personnel Management And quality management Knowledge about the evaluation of ICT systems.

Knowledge about Global economy Project management Relationship management, presentation and communication professional rules and regulations Information Technology Master Plan Security a n d safety in information and communication technology of country standard data Safety standards a n d technical standards in government information technology, rules, procedures and standards in electronic transactions.

Part 2: The part is relating to the creation of knowledge, skills and competencies as follows: Development of Systems Development Process and Data Management and Databases Design and development of programs (Programming) including testing Website design and user interaction (User Interface and Web Design) Design, manufacture and decorate animation (Animation).

Part 3: The section on application will have knowledge, skills and competencies as follows: Architecture of computers and related equipment (Hardware, Computer Architectures and Processors), Operating systems, Network and communication management. (Communications and Networks) Security Management (Network Security) Wireless and Ubiquitous Computing Service and Support

(Service Delivery and Support) such as emergency planning, building SLA or service desk, information system management Geographic Information System (GIS) Environmental Information Management (Green IT).

Part 4: General management issues will have knowledge, skills and competencies as follows: Analytical and Conceptual Thinking, Decision Making and Problem Solving, Consult (Consultation Skills), Summarizing and Writing Skills, Summarization and writing skills in English and the ASEAN community.

V. CONCLUSIONS

The research will continue in this direction, the authors start this ongoing research with the strong objectives and clear contributions. In Thailand, The structure of course syllabus according to the standard qualifications framework In accordance with the standards of the Ministry of Education of Thailand and the IEEE and ACM Computer Association. They require with a total of not less than 120 credits, consisting of general education courses Specific subjects (core subjects, elective subjects) and free elective courses In terms of specific subjects Consisting of subgroups, namely, organizational issues and information systems Technology group for applications Software technology and methods System infrastructure group. In this research, the authors have taken courses in specific subjects and elective courses as a factor in data analysis. Data mining techniques are applied to analysis. And create predictive models related to education. The authors wanted to study to develop a model for predicting the trend of getting a job that meets the qualification of information technology. Using artificial neural networks, the results obtained can be used for curriculum development. the preparation for students in the curriculum to continue to advance the information technology market with independent variables or factors that are personal information on the grades of graduates, including grades, grades in the subject, special subjects in elective subjects, foreign languages Including scores. From making a feature test that corresponds to the graduate characteristics of the information technology field and dependent variables are job positions in specialized fields of information technology.

REFERENCES
[1] Office of the Higher Education Commission, "Undergraduate standards Computer Branch," Retrieve from http://www.mua.go.th//users/he-commission/doc/law/ministry%20law/1-38%20TQF2%20computer 2552.p df, May 2016.
[2] S. Chen, B. Mulgrew, and P. M. Grant, "A clustering technique for digital communications channel equalization using radial basis function networks," IEEE Trans. on Neural Networks, vol. 4, pp. 570-578, July 1993.
[3] J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," IEEE Trans. Electron Devices, vol. ED-11, pp. 34-39, January 1959.
[4] Cooper, C.I. "A Generalizable Neural Network for Predicting Student Retention," Professional Papers Proceedings Archive, Available at: http://www.asee.org/documents/zones/zonel/2010/profession al/A-Generalizable-Neural-Network-for-Predicting-Student-R etention.pdf, 2013.

[5] Pranipa Son Ek and Suradech Bunlue. "Prediction of student achievement by using Vector Machine support," Veridian E-Journal Science and Technology Silpakorn University, 6, 40-49, 2014.

[6] Bogdan, O. "Predicting students' results in higher education using neural networks," In: Proceeding of International Conference on Applied Information and Communication Technologies (AICT2013), Jelgava, Latvia, 2013.

[7] El Moucary, C., Khair, M., Zakhem, W. "Improving Student's Performance Using Data Clustering and Neural Networks in Foreign-Language Based Higher Education," The Research Bulletin of Jordan ACM, vol II, pp. 27-34, 2011.

[8] Faculty of Information and Communication Technology, Silpakorn University, Thailand. Retrieve from https://www.ict.su.ac.th/academic/academic-it- business, May 2016.

[9] Information Technology Promotion Office Japan, "ITSS User Manual Documentation for training on ITSS / ITPE", Information Technology Skills and exam Standards(ITSS) professional qualifications, Bangkok, February 2016.

Working Principle of LI- FI Technology and its Deployment in IOT Applications

[1]**Vijaya Krishna Sonthi,** [2]**S. Nagarajan,** [3]**Koppisetti Giridhar,** [4]**M. V. B. Murali Krishna M**

[1]*Assistant Professor, Department of CSE, Sasi Institute of Technology and Engineering (A), Tadepalligudem-534101, AP., India*

[2]*Assistant Professor, Department of CSE, Government College of Engineering, Sethurappatti, Trichy-620012,TN., India*

[3,4]*Assistant Professor, Department of CSE, Sasi Institute of Technology and Engineering (A), Tadepalligudem-534101, AP., India*

E-mail: [1]*vijayakrishna1990@gmail.com,* [2]*naga72.aucse@gmail.com,* [3]*kgiridhar562@gmail.com,* [4]*muralimk786@gmail.com*

A B S T R A C T

The process of communication may happen in many ways using various means starting from expressions, body gestures, pictorial representation, letter writing and so on. Major transformation of communication was rapidly changed after invention of wired communication like using telephones. After some years the same communication process has been migrated to wireless communication with the invention of mobile phones. Meanwhile, Internet came into the picture and upgraded the communication process to next level. By using this technology, we can facilitate the transfer of information in digital format. In wireless mode of transfer, Wi-Fi technology which works on the basis of radio frequency is being used to transfer such information. But it is having some limitations and demerits regarding bandwidth, security, health issues of user, power consumption etc. The above said demerits are rectified with Li-Fi (Light Fidelity) technology. This paper will cover the basic details, design, advantages and comparison of Li-Fi technology and also discusses the usage of this technology in IoT applications.

Keywords - IoT, Wi-Fi, Li-Fi, Communication Technology

I. INTRODUCTION

The term Li-Fi (Light Fidelity) was first coined by a German Physicist named Harald Haas in 2011 at Technology, Entertainment, Design Global Talk on Visible Light Communication. The major intention of this invention is to transfer the information through illumination. Li-Fi is going to use the light source as medium to send or receive the data. Initially, this technology was designed using the intensity of LED lights. Li-Fi was invented to transfer the data faster than normal broadband. Using conventional wireless technologies, we can transfer the data up to the bandwidth of 300GHz. But, transferring through Li- Fi can support the bandwidth nearly to 400THz. Obviously, it is clear that huge amount of data can be sent using the Li-Fi technology effectively. The transmission of digital data through LED lights using VLC method can be observed with flickering according to the 1's and 0's.

If the LED is transmitting 1 in data then light is ON, and if it transmits 0 then light tends to OFF. The digital data contains consecutive 1's and 0's then the ON and OFF of LED light occurs rapidly and such flickers are not able to identify through naked eye and also makes to think the light is emitted with constant frequency. To encode the data transmitted using the light can be obtained by the rate of flickering. Using different color of LED's makes to emit data with respective frequencies. To make the transmission of Data in parallel way (sending and receiving simultaneously), an array of LED's can be used. The assumption of data transmission through Li-Fi is diagrammatically shown in the following Figure.

Figure 1. Data Transmission through Li-Fi

Internet of Things is one of the current trending technologies which makes the life sophisticated by automating the works done by manual manner. The major ideology of IoT is to make the device think and respond of their own without human involvement by communicating with each other using Internet. Such decision taking and actions performed according to the situation are carried by analyzing the data stored in cloud or server and responds to the situation using various sensors available to perform the respective tasks. IoT will use the connection protocols which utilize less power consumption and effective transmission. The devices which supports the IoT are provided with inbuilt gateways to convert the received data from various protocols to protocol supported by the corresponding device.

Using Wi-Fi technology in IoT cannot yield efficient results in all aspects because of various limitations. So, it is better to use Li-Fi technology by replacing Wi-Fi will be more beneficial in all aspects. So, the environment of using IoT with Li-Fi is illustrated in the following figure.
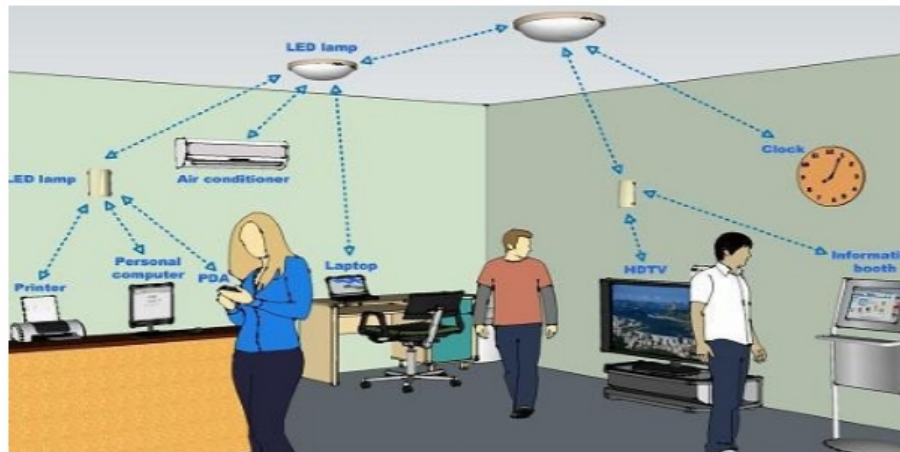
Figure 2. Environment of IoT with Li-Fi

II. DESIGN OF Li-Fi

The design of Li-Fi technology deals with architecture of Transmitter and Receiver respectively. When it comes to design of transmitter, we need to perform assemblies of different components together to perform their respective functions. Major among such assemblies is related to the architecture of Bulb. A bulb is joined with the dielectric material. This alliance can perform two functionalities which are listed below.

    a) It guides the waves of radio frequency energy sent by the power amplifier.

    b) It is used as concentrator of electric field to focus the energy into the bulb.

Such energy produces heat in the bulb to a threshold level that makes to emit the visible light with high intensity. The illuminated light intensity from the bulb varies with the color chosen in the design. Now this kind of bulb is then linked with the Printed Circuit Board which manages the electrical I/O of bulb, a Micro controller to look after various functions related to bulb. Power Amplifier is used to generate Radio Frequency signal which directs to electric field of bulb. A Receiver can be manufactured with the help of silicon photodiode which is a good recipient of the visible light. The device with above said architecture has been diagrammatically shown below.
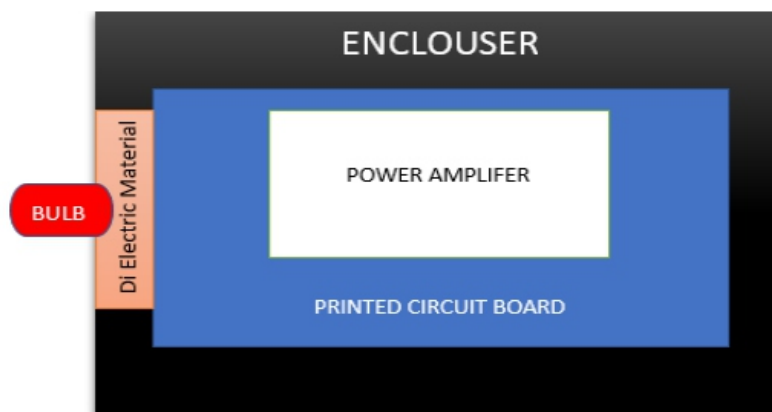
Figure 3. Architecture of Li-Fi

In order to above architectural point of view we need to consider some constraints which are to be followed in designing. They are

    a) The light should be compulsorily available for transmission of data.

    b) The Li-Fi signals will be transmitted only up to the range of light incident from a particular source.

    c) To attain better results, it is advisable to use either fluorescent or LED light sources.

By considering the above said architectural view, can helps to design the devices suitable for transmitting and receiving the data using Li-Fi technology. When the architecture of IoT has to be discussed, it consists of sensors to control and act based on the type of work to be done by the device and have connections to the server/cloud for managing the information. These connections can be done using Internet.

III. WORKING PROCESS OF Li-Fi

To use the Li-Fi technology we need a transmitter typically considered as LED light and a receiver made up of a photo detector material. The transmitter is connected to the network to modulate digital data along with the flickering light which emits from the LED light source. The receiver then receives such light and decodes the information modulated with the corresponding light rays. The working process can be shown clearly using the following figure.
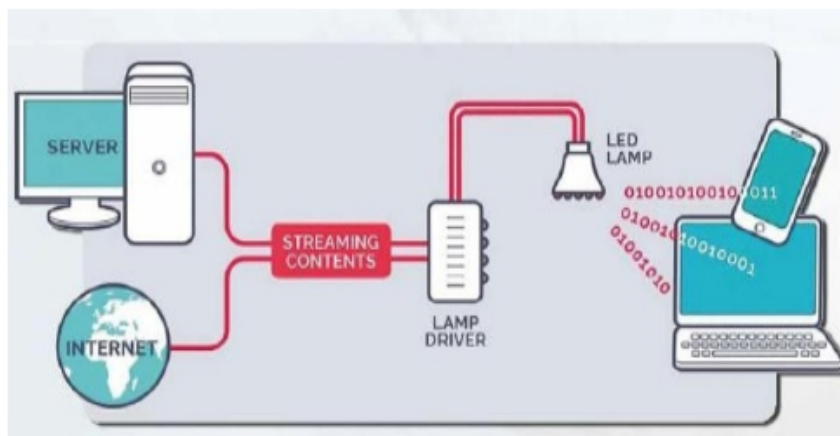
Figure 4. Working Process of Li-Fi

The major advantage of using visible light for data transmission in Li-Fi is, it doesn't have less adverse effects on the health of user. Whereas, other types of transmission mechanisms like Bluetooth, NFC, Wi-Fi will use different frequencies like X-ray, Infrared Rays, Gamma Rays, Radio Frequencies etc. which can show harmful effects on user's health.

The biggest reason of using visible light communication (VLC) is to highlight the systems that supports middle level communications performs at less speed like mobile to mobile and fixed to mobile communications which transfers data at high rate. Data rates which uses 1 Gbps range for different modulation schemes. IEEE 802.15.7 defines physical layer (PHY) & media access control (MAC) layer for VLC/Li-Fi. The MAC layer supports 3 multi-access technologies: peer-to-peer, star configuration and broadcast mode. It deals physical layer management problems like addressing, collision avoidance and data acknowledgement protocols. The physical layer is classified into 3 types.

- The PHY I was established for outdoor application and works from 11.67 kbps to 267.6 kbps.
- The PHY II permits reaching data rates from 1.25 Mbit/s to 96 Mbit/s.
- The PHY III is used for many emissions sources with a particular modulation method called colour shift keying (CSK) and data rates may vary from 12 Mbit/s to 96 Mbit/s.

IV. MERITS OF Li-Fi AND IoT

Utilizing the visible light communication (VLC) for data transmission we can have several advantages over remaining wireless transmission techniques. The major part used in this kind of transmission is happen with an LED bulb as light source, it makes cheaper to incorporate as per the needs, also helps to lessen the consumption of power. By joining high intensity light sources along with higher bandwidth, such source can support more than the data rate of 1Gbps. Li-Fi is used within the bounded limit and cannot be intervened or misused by others.

IoT have various uses in different aspects. For Example: enormous data is available for sensors to take necessary actions according to the situation, tracking of various things are automatically done by the devices itself can help the user to utilize the amount of time taken to track such activities for useful work, these kinds of activities will reduce the involvement of man power.

V. BARRIERS OF Li-Fi AND IoT

Barriers of Li-Fi:
- Absence of light source doesn't allow the Internet access.
- Needs consistent Line-of-Source for data transmission.
- Meanwhile of transmission, any of existing transparent objects will obstruct the data which is being transmitted.
- Range of Li-Fi is short when compared with Wi-Fi and cannot able to move through hard surfaces.
- For complete establishment of this technical network, installation cost will be very high.
- It has a long way to reach the common people as compared with another wireless technologies.

Barriers of IoT:

- Common concept accessed and used by all devices like USB is not yet developed for IoT.
- Propagation of a single message for multiple times to the users connected for a single device may happen.
- Encryption of the utilised data is not possible in this technique makes others to get details of a user is easier. So, it has high risk of data theft.

VI. APPLICATIONS

- The combination of Li-Fi and IoT can be used in education system to transmit/receive the data among various devices in institutional level. By making this to come true, all the educational institutions will be benefited with secure and uninterrupted data transmission with cost effective measures.
- In medical field, especially for surgeries this combination is the best so far. Because, to transmit the data between devices Wi-Fi usage is not preferable as it transmits through RF waves it may affect the health of patient. So, to transmit the data we use VLC and also IoT to regulate the devices in the room on their own based on the information gathered by the sensors and analyses based on the stored and collected data, acts as per the situation.
- The combination of IoT and Li-Fi can be used in both aircraft and underwater vehicles like Aircraft, Submarines etc as data transmission for crew and passengers, devices to communicate themselves. The main intension to use this combination in underwater vehicles can help for data transmission where the existing Wi-Fi technology completely fails to serve data using wireless mode.
- It can be used in the stream of disaster management to reduce the human loss by transmitting the rescue related information, emergency alerts by various departments using the Li-Fi which is very efficient and easy to reach the common people.
- This technology can be used extensively in the sensitive locations like power plants, petrol bunks, and other restricted areas where the Wi-Fi was restricted.
- Connecting the nearby devices like mobiles, laptop, camera, TV etc. can be well handled for short range networks with high security and data rate.
- In the field related to health IoT embedded with Li-Fi can bring drastic changes in daily life. By using this we can emerge the kits used to check minimum requirements of patients like Heart rate, Blood Pressure, Diabetic reports, and also, we can provide the medication support by analysing the data like previous cases, basic symptoms, preventive methods, supportive medicines, best doctors in nearby areas so on. These things can happen with the help of combined technologies like GPS, WAN etc.

It will help farmers to find the fertility of soil, available water resources, analysing the nature of climate and soil, suggesting favourable crops which can get profits with less income, fertilizers helpful for the farms etc.

We use these technologies to regulate the traffic in congested roads by gathering and suggesting the alternative routes in busy times, nearest vacant parking slots, road conditions etc.

VII. CONCLUSION

This paper gives the information regarding the functionality of Li-Fi technology and emerging trends related to Internet of Things. Also explains various advantages and disadvantages of respective technologies. Various applications are listed which can be used by either IoT or Li-Fi. The comparisons among standard wireless transmission technique like Wi-Fi and Li-Fi is also discussed for better understanding of the advents of Li-Fi. Even though the discussed technology can have some drawbacks, but it is considered as the most reliable transmission technique so far as per the health issues of users. It doesn't harm the users through its medium and type of transmission frequency.

When comes to IoT, it is the recent technology which is extensively using in the process of automation. The major functionality and usage of this technique is explained and suggested some of the advantages in application orientation. The respective technique will transmit the data using regular mode of wireless standard called Wi-Fi. And the sensors attached with the devices can perform according to the operations done by retrieving the existing data stored in the servers or cloud for analyzing and to act as per the situation without human intervention.

FUTURE SCOPE

As per the studies, it is theoretically proven that Li-Fi can show better performance than Wi-Fi. And Wi-Fi is being used by the devices to support IoT mechanism, such devices can be upgraded to the Li-Fi technology instead of Wi-Fi can generate high productive results in all aspects with rapid speed, less cost high security, effective utilization.

REFERENCES
[1] P.P. Ray, 2018, "A survey on Internet of Things architectures", Journal of King Saud University – Computer and Information Sciences, 30, 291–319
[2] Mostafa H. Ahfayd, Zohaib A. Farhat, Martin J. N. Sibley, Peter J. Mather, and P. I. Lazaridis, 2018, " Selection of high power LEDs for Li-Fi applications", IEEE 25th conference on ICT, pp 170-174
[3] Haesung Lee, Kwangyoung Kim and Joonhee Kwon, 2016, "A Pervasive Interconnection Technique for Efficient Information Sharing in Social IoT Environment", International Journal of Smart Home Vol. 10, No. 1, pp. 9- 22.
[4] Kajal, Ashish Saini Kanchan Gupta, "Lifi- Light Fidelity Technology- A Review," International Journal of Emerging Research in Management &Technology, vol. 3, no. 10, Oct 2014.

[5] M. Mutthamma, "A survey on Transmission of data through illumination - Li-Fi," International Journal of Research in Computer and, Dec 2013.

[6] Jyoti Rani, Prerna Chauhan, Ritika Tripathi, 2012, "Li-Fi (Light Fidelity)-The future technology In Wireless communication", International Journal of Applied Engineering Research, ISSN 0973-4562 Vol.7 No.11.

[7] Lamya I. Albraheem , Lamia H. Alhudaithy, Afnan A. Aljaser, Muneerah R. Aldhafian, Ghada M. Bahliwah, 2018, " Toward Designing a Li-Fi-Based Hierarchical IoT Architecture", IEEE Access Volume 6, pp 40811 - 40825.

[8] Richard Gilliard, "THE LiFi® LAMP HIGH EFFICIENCY HIGH BRIGHTNESS LIGHT EMITTING PLASMA WITH LONG LIFE AND EXCELLENT COLOR QUALITY," IEEE, Sunnyvale, USA, 2010.

[9] http://purelifi.com/

[10] http://www.ted.com/talks/harald_haas_wireless_data_from_every_light_bulb

[11] https://www.linkedin.com/pulse/what-lifi-technology-how- does-work-beginners-guide-akash

[12] https://www.led-professional.com/resources-1/articles/lifi-what-it-is-how-it-works-what-it-provides-how -to-apply-and- its-future-prospects

Building and Implementation of Naive Bayes Classification Model: A Domain of Educational Data Mining

[1] Abdullah Saadal- Malaise Al- Ghamdi, [2] Farrukh Saleem

[1,2]Information Systems Department,Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Kingdom of Saudi Arabia,
E-mail: [1] aalmalaise@kau.edu.sa, [2] fsaleem@kau.edu.sa

A B S T R A C T

Data mining is a technique uses to identify and extract the hidden knowledge from the dataset. It has vast range of applications such as hospital, weather forecast, business industries, and others. Now a days, Educational Data Mining (EDM) is one of the common applications use for analyzing the data related to students, course, and faculty performance. This paper discusses the EDM using Naïve Bayes classification algorithm. The purpose of this research is to identify the relative pattern in the collected dataset from educational database to predict the performance of the students. Therefore, a conceptual framework presented to understand the main theme of this research and execution of the experiment. The framework implementation applied using Rapid Miner tool, whereas the model presented in this research indicates the preparation of training and testing dataset to predict the student's potential performance. The result elaborates the positive execution of the model, which is being proved by the "performance" operator provided in rapid miner tool. The accuracy of model measured as 96.43%, that highlights the encouraging impact of the EDM model presented in this study. The research can be useful for the college/university instructors to predict the performance of their students in ongoing and future semesters.

Keywords - Educational Data Mining; Classification; Naïve Bayes Algorithm; Prediction.

I. INTRODUCTION

The practice of advanced technologies and moderntools implemented in higher educational institutes created large amount of data everyday. Proper management and transformation of these into useable formats and procedures can enhance the constructive influence of stored data onto organization. The use of data in order to search hidden patterns and rules can further improve the organizational business processes as well. Data mining is the tool, which can help the organization to identify and extract valuable and hidden information from the database and historical data warehouse [1]. The data mining techniques provide analytical and statistical approach, which further can elaborate using visualization tool to highlights the interesting facts extracted from the data set. The data collected from any sources can be useful to apply any data mining techniques such as classification, association or clustering [2]. It is crucial to understand data before applying any techniques as data mining algorithm developed based on the context of the stored data. For example, classification algorithm's approach is based on supervised learning where a label or class attribute is necessary to build and validate the model [3]. On the other side, to generate new rules from the data by learning and associating the selected attributes from the previous data file, the association mining algorithms are useful [4]–[7]. EDM is one of major

applications of data mining presented in different context. For example, predicting student's performance using classification algorithm, where the results highlighted the encouraging points from lecturer's and students perspectives [8]. Another research presented by [9] to understand the performance of student and predict their passing ratio in engineering major. The program used in this research was based on Java, whereas Naïve Bayes algorithm used to forecast the student's performance. In addition, building academic trends and patterns using educational database [10], the work presented by implementation of association mining technique using Apriori algorithm. The research generated rules and patterns based on several students' records such as attendance, practical exam, projects, student's profiles and others. [11] described the complete literature in order to generate rules and patterns from any information system or application, the theory presented in this research can be useful for EDM as well. This study focused on predicting the student's performance in the upcoming semesters using their performances and participation collected from previous semester(s). The main problem highlighted in this research is how to generate positive impact from student's historical data? Furthermore, this research also delivers the comprehensive guidelines for the instructors to understand the potential performance in their respective courses. Therefore, based on the literature review collected, the Naïve Bayes algorithm implemented using Rapid Miner tool. The data analysis techniques, list of attributes, and data values are discussed in later section.

II. EDUCATIONAL DATA MINING USING RAPID MINER

Rapid miner is the tool used in this research for the proposed model. It is an open source tool use for the implementation of Classification algorithm; Naïve Bayes. The rapid miner tool is widely used by research and educational community. It provides the list of operators to perform required activity such as data preprocessing, modeling, predication and machine learning. The interface is user friendly, while on each step guidelines is provided by rapid miner team[12]. It further provides an integrated development environment to build and apply data mining tasks with comprehensive list of operators and example data sets for the support of researchers and students[13].

There are several researches found shown the successful implementation of educational models and experiment using rapid miner. For example, the implementation of data mining algorithm on student data [14], the reviews on prediction algorithm in EDM in rapid miner[15]. Another research described the middle-school module for Introducing Data- Mining using rapid miner [16]. It highlights the rapid miner's usefulness for EDM or other applications.

III. PROPOSED MODEL: IMPLEMENTATION AND VALIDATION

The process execution performed using the framework presented in Fig.1. The model describes, as the first step is to collect data from educational database. Data preprocessing is the further step to be

accomplished in order to remove noisy data, normalize the data, and to transform it in suitable format. In the next step, data split into training and testing dataset to perform building and testing of the model using Naïve Bayes algorithm. Furthermore, the classification algorithm applied on preprocessed data using Rapid Miner. The subsequent sections describes the description of algorithm used in this research, data attributes and analysis phase, finally result and discussion presented to understand the main objectives of this research study.
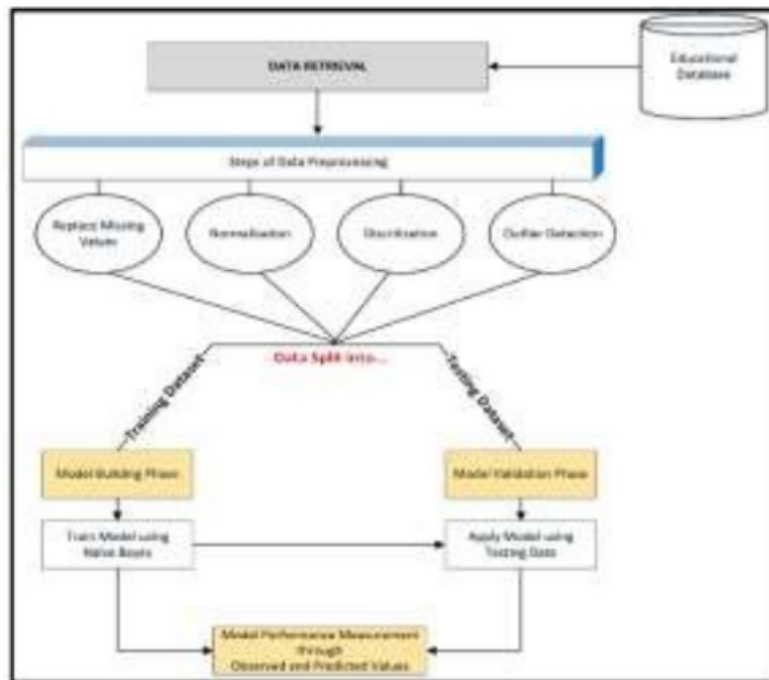
Fig.1. Framework for Implementation of Student's Performance and Measurement Model

Naïve BayesAlgorithm

The Naïve Bayes is the classification algorithm most commonly used for extraction new information from selected data. This algorithm uses unseen data during training phase, which must have a class column. The Naïve Bayes classifier assigns each data object as a class based on highest conditional probability [3]. Apart from its simplicity in implementation, the Bayesian classifier interestingly perform well in learning from the given dataset and to forecast for future transactions [9]. The developed model can further use for validation purposes to predict the class for the unlabeled data during testing phase. The pseudocode of Naïve Bayes algorithm is presented below in the Fig.2.

```
Input:
        Training dataset T,
        F= (f₁, f₂, f₃,... fₙ)    // value of the predictor variable
        in testing dataset.
Output:
        A class of testing dataset.
Step:
    1.  Read the training dataset T;
    2.  Calculate the mean and standard deviation of the
        predictor variables in each class;
    3.  Repeat
            Calculate the probability of fᵢ using the gauss
            density equation in each class;

            Until the probability of all predictor variables (f₁, f₂,
            f₃,... fₙ) has been calculated.
    4.  Calculate the likelihood for each class;
```

Fig.2. Pseudocode of Naïve Bayes algorithm[17]

Data Analysis and Preprocessing

In this study, the educational database used for collecting and analyzing the data to predict the performance of the students. Data collecting from the various sources, where the students were examined based on several assessment methods, such as assignment, quizzes, exam, projects, group discussions, behavior, and participation. The list of the attributes selected in this experiment is illustrated in Table1. All of the attributes were selected for training and testing phase, while the performance predicted based on these attributes, as defined in next section.

Attributes	Description	Values
Assignments	The marks given in different assignments, the average value is selected for the experiment.	between 1 to 10
Quizzes	The marks given in different quizzes all over the semester, the average value is selected.	between 1 to 10
Exams	There are different exam taken i.e. mid and final exam. The average of all exams computed.	between 1 to 50
Group Project	As group project marks are given as group-wise. The marks has been divided to each member of the group.	between 1 to 20
Participation	Student's participation in the class while performing different activities	Average Good
		Excellent
Student's behavior	Overall behavior of students from learning and dealing perspectives	Average Good
		Excellent
Attendance	Attending lectures and other activities.	Low Average
		High
Overall Performance	This attribute describes the overall performance measured in the semester. This attribute is also	Average Good
	considered as a label/class column.	Excellent

Table1: List of Attributes for Experiment

After selection of the data, the data preprocessing steps applied on dataset to prepare and transform it in an adequate manner as this is an important step in data mining [4]. Stepwise approach applied on the data for experiment preparation purpose. Initially, 760 students' records were selected, later on 56 records were removed due to duplicate entries in some attributes. In addition, 49 records were eliminated as found some attributes values were missing. Furthermore, some attributes values transformed into particular type so that they can be appropriate for Naïve Bayes algorithm implementation. Finally, to perform the experiment total 655 records were selected for the training and testing phase. The detailed implementation of preprocessing steps using Rapid Minerare shown in Fig.3.
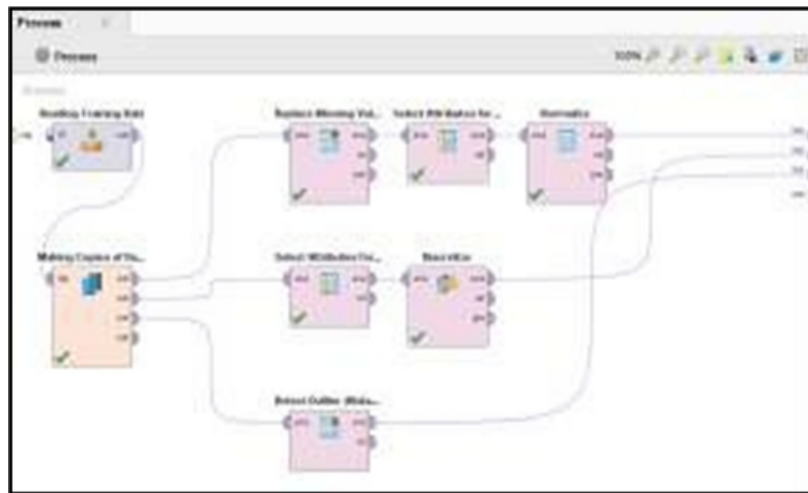
Fig.3. Data Preprocessing using Rapid Miner

Model Implementation and Validation Applying classification algorithm and building Bayesian classifier, the training phase plays an important role. The purpose of training phase is to identify the well-known patterns in the data set and to build an accurate model, which can exemplify the real world problem in an efficient manner [18]. The preprocessed data selected for training phase, where as all the attributes written in Table-1 were part of training phase. As part of the classification training phase, one of the attributes must be selected as a "Class/Label" column, therefore "Overall Performance" is considered as Label or Class column. This column will help to train the model and predict the performance for upcoming semester. Basically, the assumption of selecting the label (class) and its value is use as calculating conditional probability based on the given other attribute for the same row[19]. This hypothesis not always right but it's naïve model, which shows simplification in measuring the probability. As found in previous studies, which highlighted that the naïve assumptions successful ratio is good enough to predict the academic performance [8], [9]. The sample of data values under each variable is shown in Fig.4.

S.No	Student	Assignments	Quizzes	Exams	Group Project	Participation	Student's Behavior	Attendance	Overall Performance [Label Column]
1	Female	1	1	8	8	Average	Good	Average	Average
2	Female	9	9	5	11	Good	Average	High	Excellent
3	Male	8	9	32	14	Good	Good	Average	Good
4	Female	1	8	21	18	Excellent	Good	High	Excellent
5	Male	1	8	41	1	Average	Average	Low	Average
6	Female	8	8	31	19	Excellent	Good	High	Excellent
7	Male	1	4	25	8	Average	Average	Low	Average
8	Female	4	2	21	8	Average	Excellent	Average	Good
9	Female	6	7	1	19	Average	Average	Low	Average
10	Male	7	1	15	12	Excellent	Average	High	Good
11	Male	3	4	21	14	Average	Excellent	Average	Good
12	Male	9	1	16	11	Average	Average	Low	Average
13	Female	4	10	48	20	Excellent	Good	High	Excellent
14	Female	2	10	16	18	Average	Average	Average	Average
15	Male	18	4	43	14	Average	Excellent	Low	Average
16	Male	6	5	22	14	Excellent	Good	High	Excellent
17	Male	8	8	49	9	Average	Average	Average	Average
18	Female	9	1	8	18	Average	Average	Low	Average

Fig.4. Sample of Dataset and Attributes

The process implementation using Naïve Bayes algorithm is applied on selected data set. The model is developed in rapid miner to run this experiment. To run the model, there are several operators need to be connected in particular sequence to execute it successfully. The description of each operator is shown in Table 2.

S.No.	Operator Name	Description
1	Data Set (Read CSV)	This Operator reads an ExampleSet from the specified CSV file.
2	Split Data	This operator produces the desired number of subsets of the given ExampleSet. The ExampleSetis partitioned into subsets according to the specified relative sizes.
3	Naïve Bayes	This operator generates a Naive Bayes classification model.
4	Apply Model	This Operator applies a model on an ExampleSet.
5	Performance	This operator is used for statistical performance evaluation of classification tasks. This operator delivers a list of performance criteria values of the classification task.

Table2: Description of Operators Used in the Experiments[12]

The model developed in rapid miner starts with the selection of data set. The rapid miner tool allows the user to link the data or copy the selected data in rapid miner's own repository. The stepwise approach of this experiment is shown below, whereas Fig.5 highlights all operators and their connectivity in rapid miner.

- Step-1: The selected data set is connected with the operator called "Data Set-(Read CSV)".
- Step-2: The Data file has been split into two parts as training data and testing data, as 70% and 30% respectively.

- Step-3: On training dataset (70% of the whole data), Naïve Bayes algorithm applied to train the model. The model works as measuring the conditional probability between "class column" and other attributes.

- Step-4: The trained model further sent to the operator called "Apply Model". The purpose is to validate (test) the model using remaining data set known as Testing Data (30% of the whole data) selected randomly.

- Step-5: Finally the performance measurement operator used to measure the accuracy of the model.
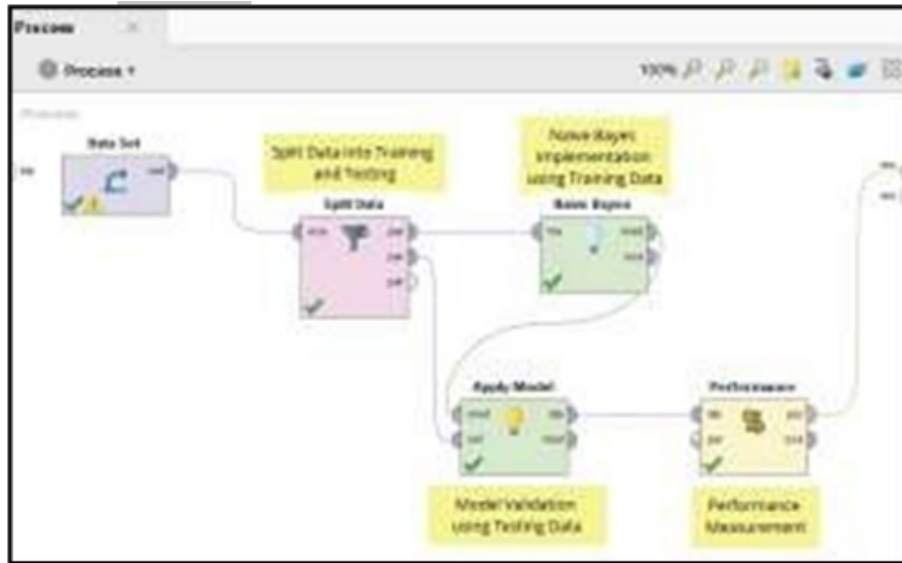
Fig.5. Implementation of the Proposed Model Using RapidMiner

RESULTS AND DISCUSSION

The vast amount of data stored in educational databases, the proper use of the data can provide valuable results, which can further create positive impact on educational environment and performance of the students. This experiment highlighted the appropriate use of educational dataset where students and teacher can take the benefits as equally. The classification algorithm used to measure the conditional probability using Bayesian classifier in order to learn the student's performance from previous data and to predict their performance in future. Bayesian classifier works by studying and observing the previous data provided as training data. In this experiment, 70% of the data was selected as training data, which were exactly 459 records of student. The sample of dataset shown in Fig.4. The experiment further implemented based on the conceptual framework presented in this study as shown in Fig.1. The conceptual framework is executed using rapid miner with the help of operators connected with each other as per the requirements. The results indicates that the validation of the model is achieved successfully with 96.43% accuracy. The accuracy of the model is shown in Fig.6 in the form of confusion matrix provided by result section of rapid miner. According to the training phase (459 instances) in this experiment, the prediction has been recorded in the form of "Average", "Good", and

"Excellent". The class of student's performance was pre-defined in the training data set, as the experiment used supervised learning approach. In supervised learning approach the categories of classes are pre- defined depending on other list of attributes. Therefore, the testing or validation phase in this study only describes the performance according to the distinct values, which already defined in label or class column during training phase. The validation phase use the testing data that was 30% of the whole data selected in this study. In division, there were total 196 instances selected for testing or validation phase. Out of those 196 instances, the accuracy results as shown in Fig.6, illustrates that, 103 instances belong to the student whose performance was "Average" and all of them predicted as average, which demonstrates the 100% of prediction for the performance of average students. Same100% accuracy recorded for the "Excellent" students. Although, the students belong to the category as "Good" student, the class precision for those students measured as 88.33%. There were total 60 instances belongs to this category, in which 53 predicted accurately, whereas 7 students prediction was wrongly done. Therefore, the overall performance of the model is measured as 96.43%.

Fig.6. Accuracy of the Model Using Confusion Matrix

In addition, the following Fig.7, showing the histogram of the label column known as "Overall Performance". This figure is highlighting the frequency of each distinct values predicted after validation of the model under class column. For example, for "Average" student it is 103, for "Excellent" it is 33, and finally for "Good" category, the number of students were 60.
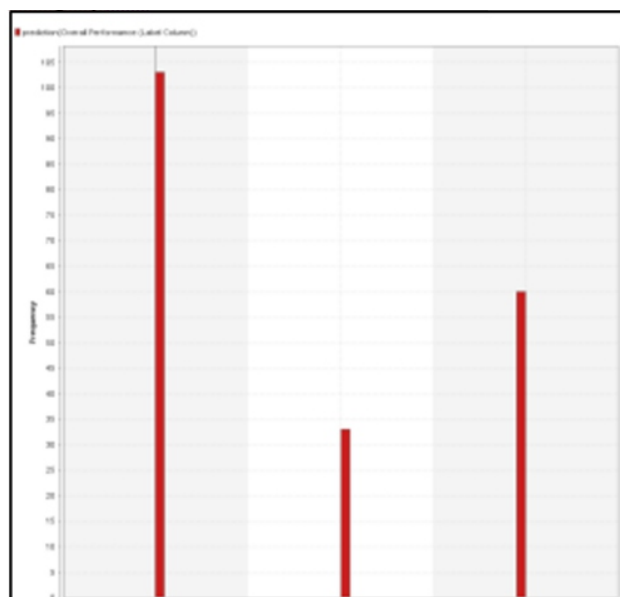

Fig.7. Histogram of Overall Performance

IV. CONCLUSION

Data mining techniques are useful to understand and extract the hidden knowledge from the database. In this study, the EDM domain selected for the implementation of data mining classification algorithm. For this, the Naïve Bayes method selected to predict the performance of the student based on the previous grading scales and criteria. Firstly, this study presented the conceptual framework to provide the clear guidance for the readers. The model then implemented using Rapid Miner; a data mining and machine-learning tool. For model building, the data extracted from the educational database, later on the data split into training and testing phase to build and test model. The results obtained through the experiments measured the 96.43% of the model accuracy, which illustrate the quite fair results for educational data mining purposes. The results can be helpful for the instructors in order to understand the student's capabilities and performance in upcoming semesters. In future, the conceptual model and experiment can be applied on other databases, while comparison between different classification algorithms can be presented in future.

REFERENCES

[1] F. Saleem and A. AL-Malaise AL-Ghamdi, "Implementation of data mining approach for building automated decision support systems," in International Conference on Information Society, i-Society 2012, 2012, pp. 127–130.

[2] F. Saleem and A. Malibari, "Data mining course in information system department- case study of King Abdulaziz University," in 2011 3rd International Congress on Engineering Education: Rethinking Engineering Education, The Way Forward, ICEED 2011, 2011.

[3] A. Dangi and S. Srivastava, "Educational data Classification using Selective Naïve Bayes for Quota categorization," in 2014 IEEE International Conference on MOOC, Innovation and Technology in Education (MITE), 2014, pp. 118–121.

[4] A. S. Al-Mudimigh, F. Saleem, and Z. Ullah, "The effects of Data Mining in ERP-CRM model - A case study of MADAR," WSEAS Trans. Comput., vol. 8, no. 5, 2009.

[5] A. S. Al-Mudimigh, Z. Ullah, F. Saleem, and F. N. Al- Aboud, "Data mining for customer queries in ERP model - A case study," in NCM 2009 - 5th International Joint Conference on INC, IMS, and IDC, 2009.

[6] A. S. Al-Mudimigh, Z. Ullah, and F. Saleem, "Data mining strategies and techniques for CRM systems," in 2009 IEEE International Conference on System of Systems Engineering, SoSE 2009, 2009.

[7] A. S. Al-Mudimigh, F. Saleem, Z. Ullah, and F. N. Al- Aboud, "Implementation of Data Mining Engine on CRM - Improve customer satisfaction," in 2009 International Conference on Information and Communication Technologies, ICICT 2009, 2009.

[8] F. Razaque, N. Soomro, S. Ahmed, S. Soomro, J. A. Samo, and H. Dharejo, "Using Naïve Bayes Algorithm to Students 'bachelor Academic Performances Analysis," in Engineering Technologies and Applied Sciences (ICETAS), 2017 4th IEEE International Conference, 2017, pp. 1–5.

[9] V. Patil, S. Suryawanshi, M. Saner, V. Patil, and B. Sarode, "STUDENT PERFORMANCE PREDICTION USING CLASSIFICATION DATA MINING TECHNIQUES," Int. J. Sci. Dev. Res., vol. 2, no. 6, pp. 163–167, 2017.

[10] F. Parack, S., Zahid, Z., & Merchant, "Application of data mining in educational databases for predicting academic trends and patterns," in IEEE International Conference on Technology Enhanced Education, 2012, pp. 1–4.

[11] Y. L. Wang, P. J., Shi, L., Bai, J. N., & Zhao, "Mining association rules based on Apriori algorithm and application," in International Forum on Computer Science-Technology and Applications, 2009, pp. 141–143.

[12] R. M. Team, "Rapid Miner." [Online]. Available: https://rapidminer.com/.

[13] F. Saleem and A. Malibari, "DATA MINING COURSE IN INFORMATION SYSTEM DEPARTMENT–CASE STUDY OF KING ABDULAZIZ UNIVERSITY," in 3rd International Congress on Engineering Education, 2011.

[14] S. Angra and S. Ahuja, "Implementation of Data Mining Algorithms on Student's Data using Rapid Miner," in International Conference On Big Data Analytics and computational Intelligence (ICBDACI), 2017, pp. 387–391.

[15] K. S. Kumar, A. D., Selvam, R. P., & Kumar, "Review on prediction algorithms in educational data mining," Int. J. Pure Appl. Math., vol. 118, no. 8, pp. 531–537, 2018.

[16] A. Dryer, A., Walia, N., & Chattopadhyay, "A Middle- School Module for Introducing Data-Mining, Big-Data, Ethics and Privacy Using RapidMiner and a Hollywood Theme.," in In Proceedings of the 49th ACM Technical Symposium on Computer Science Education, 2018, pp. 753– 758.

[17] T. Saputra, M. F. A., Widiyaningtyas and A. P. Wibawa, "Illiteracy Classification Using K Means-Naïve Bayes Algorithm," Int. J. Informatics Vis., vol. 2, no. 3, pp. 153– 158, 2018.

[18] T. Y. Xie, X., Ho, J. W., Murphy, C., Kaiser, G., Xu, B., & Chen, "Testing and validating machine learning classifiers by metamorphic testing," J. Syst. Softw., vol. 84, no. 4, pp. 544– 558, 2011.

[19] RapidMinerTeam, "Naive Bayes Documentation," Version 9.2. [Online]. Available: https://docs.rapid miner.com/latest/studio/operators/modeling/ predictive/bayesian/naive_bayes.html. [Accessed: 15-Feb-2019].

A Set of Convolutional Neural Networks for Person Identification with Different Biometrics

[1] Goutam Sarker, [2] Swagata Ghosh

[1]SMIEEE, Associate Professor, CSE Dept., NIT Durgapur, India
[2]M.TECH. Student, CSE Dept., NIT Durgapur, India
E-mail: [1]g_sarker@ieee.org, [2]ghosh.swagata6@gmail.com

A B S T R A C T

Biometric Identification is one of the most intriguing fields of research in the broad area of pattern recognition and computer vision and it finds wide applications in government institutions, forensic research and crime investigation, employee and student attendance roll, restricted access to classified information, and in various public as well as private enterprises. Human face is an age-old biometric and recognition is based on the unique facial structure and symmetry of an individual. Fingerprint is a very important biometric for forensic research as it can be collected easily from the crime scene and also every individual has unique patterns on every finger. Human iris has unique patterns for any particular individual and thus iris verification is gaining ground as an important person identity authentication method. In the present paper these three biometrics have been chosen because these are commonly available and most frequently used. The present Convolution Neural Network (CNN) based person identification system with different biometrics is simple but efficient, effective, and fast. The performance evaluation of the system in terms of accuracy, precision, recall and f-score derived from confusion matrix together with training and testing time is moderate, appreciable and affordable.

Keywords - Pattern Recognition; Convolution Neural Network; Biometric; Holdout Method; Confusion Matrix; Face; Fingerprint; Iris; Accuracy; Precision; Recall; F-Score.

I. INTRODUCTION

Biometric identification is a security process that relies on the unique biological characteristics of an individual to verify that he/she really is the individual who he/she claims to be. Biometric identification systems compare the biometric data captured to the learnt authentic data in a database. The most commonly available and used biometrics are face, fingerprint and iris. CNN is a revolutionary and dramatic change in the field of neural networks, and it has helped computer scientists achieve astounding improvement in lowering the error rates in pattern recognition, image classification and other applications in the field of computer vision. In the present paper the three different biometrics namely face, fingerprint and iris have been used in the CNN model for person identification. The three different CNN models are trained using three different biometrics namely face, fingerprint and iris of different persons.

II. RELATED WORKS

Previously several biometric authentication systems have been implemented using classical neural networks. Hurieh Khalajzadeh et al. have come up with a hybrid system that combines a CNN to detect

and recognize facial images and a Logistic Regression Classifier (LRC) to classify the features learned by the CNN.[19] Huiying Hu et al. have developed two CNN models namely CNN-1 with two convolution layers and CNN-2 with one convolution layer for testing on 2D and 3D facial image dataset.[18] Zhonghua Liu et al. have presented a method for face recognition that combines wavelet transformation with singular value decomposition and kernel principal component analysis.[20] S. Kundu and G. Sarker have worked on a Modified SOM- Based RBFN for Rotation Invariant Clear and Occluded Fingerprint Recognition.[9, 7] D. Bhakta and G. Sarker have worked on Location and Rotation Invariant Face Identification and Localization with or without Occlusion using Modified RBF Network.[12, 14] Again, S. Kundu and G. Sarker have also worked on A Modified BP Network using Malsburg Learning for Rotation and Location Invariant Fingerprint Recognition and Localization with and without Occlusion.[16, 17, 14, 10, 11]Their research in this area mainly focuses on four categories – same classifier for the same biometric, same classifier for different biometrics, different classifiers for the same biometric and different classifiers for different biometrics. The present work showcases an example of the same CNN classifier for different biometrics.

III. OVERVIEW OF THE SYSTEM AND ALGORITHMS

The present section presents the details of the convolutional neural network architecture that is used to implement the application of the CNN-based Person Identification System with different biometrics namely face, fingerprint and iris.

1. CONVOLUTIONAL NEURAL NETWORK (CNN)

Ordinarily a CNN is comprised of three different layers namely convolutional layer, pooling layer and fully-connected layer. When these layers are stacked together, a complete CNN architecture is formed. [5, 6]

The convolution layer is used to produce the activation map for all the features by convolving or sliding a kernel or filter across every location of the pixel matrix of the input image. This layer is essential to find out which feature exists especially in which part of the image. [5, 6]

The convolutional layer of the CNN determines the output of neurons connected to local regions of input through the calculation of the scalar product between their weights and the regions connected to the input volume. The rectified linear unit (abbreviated as ReLU) is used to apply an 'elementwise' activation function. The ReLU function is more advantageous than the sigmoid activation since it decreases the possibility of vanishing gradient where the weight updating process slows down significantly. [5, 6]

The pooling layer of the CNN simply performs down sampling along the spatial dimensionality of the given input, further reducing the number of parameters within that activation. [5, 6]

Finally, after several convolution and pooling layers, the fully-connected layer takes care of the high-level reasoning of the neural network by performing the same duties as in standard Artificial Neural Network (ANN), and the neurons in this layer have connections to all the activations. [5, 6]

2. THEORY OF OPERATION

Pre-processing of Face, Fingerprint and Iris Images

The images of the three different biometrics namely face, fingerprint and iris have to be pre-processed before CNN learning as well as recognition. The required pre-processing steps are detailed below:

i) Conversion of RGB images to grayscale images: The first step of preprocessing is to convert training or test image set into gray scale image set. Those images may or may not be noisy.

ii) Removal of noise from the images: The images may be noisy. The second step is to remove those noises from face patterns of the databases. In this process we use Dilation followed by Erosion methods.

iii) Conversion of gray-scale images into ID matrix: In the last step of preprocessing, 2D matrix image files are converted into 1D matrix files. This set is the input to the convolutional neural network for learning and recognition.
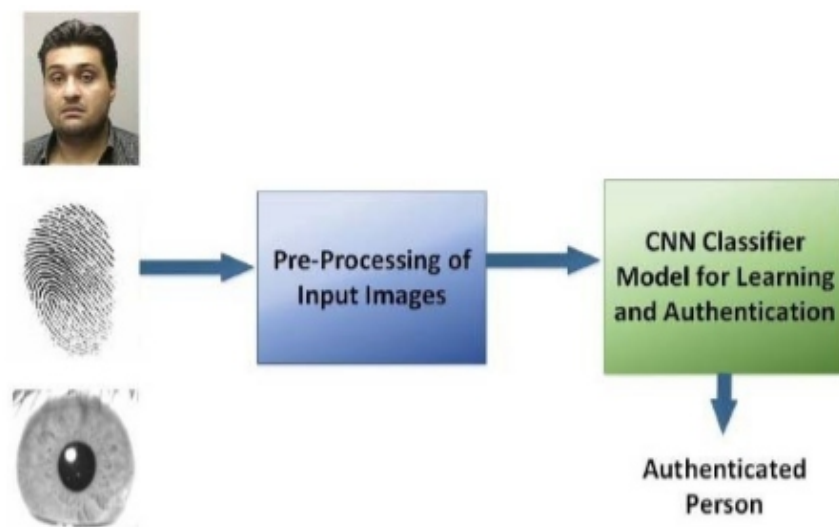
Figure 5. Overview of the System

Operation of the CNN

The network has input image of size 128*128 neurons. These are the pixel intensities for the biometric data set images. Then a convolution layer with 50 kernels each of size 3*3 (this is also the size of the local receptive field) follows this. This results in 50*126*126 hidden feature neurons. The next step is to perform max pooling with a max pooling kernel size 2*2 with a stride of 2 (no overlapping) across all the 50 feature or activation maps. The result is a layer of 50*63*63 hidden feature neurons. The process of convolution and max pooling is once again repeated, with 50 kernels of size as before i.e. 3*3. With this, convolution layer yields a layer of 50*61*61 hidden feature neurons, while the max pooling layer yields a layer of 50*30*30 hidden feature neurons. This is now flattened at the next layer. [5, 6]

The final layer of connection is a fully connected layer. So this layer connects every neuron from the flattened max pooled layer to every one of the 10 output neurons (as there are 10 categories for classification), in the initial learning stage for each case of facial image CNN, fingerprint CNN and iris CNN; and in the latter stage of CNN learning there are 25 output neurons for 25 classes of the face CNN, 40 output neurons for the 40 classes of fingerprint CNN and 36 output neurons for the 36 classes of iris CNN. [5, 6]

Figure 6. Step-by-Step Approach in Keras CNN Model

Algorithm for Training the CNN for Recognizing Face, Fingerprint and Iris Images Input: Set of pre-processed images of face, fingerprint and iris each of size 128x128 pixels

Output: Three CNN models that are trained separately for the identification of face, iris and fingerprint

Steps for training each CNN model:

1. The network has input image of size 128*128 neurons.
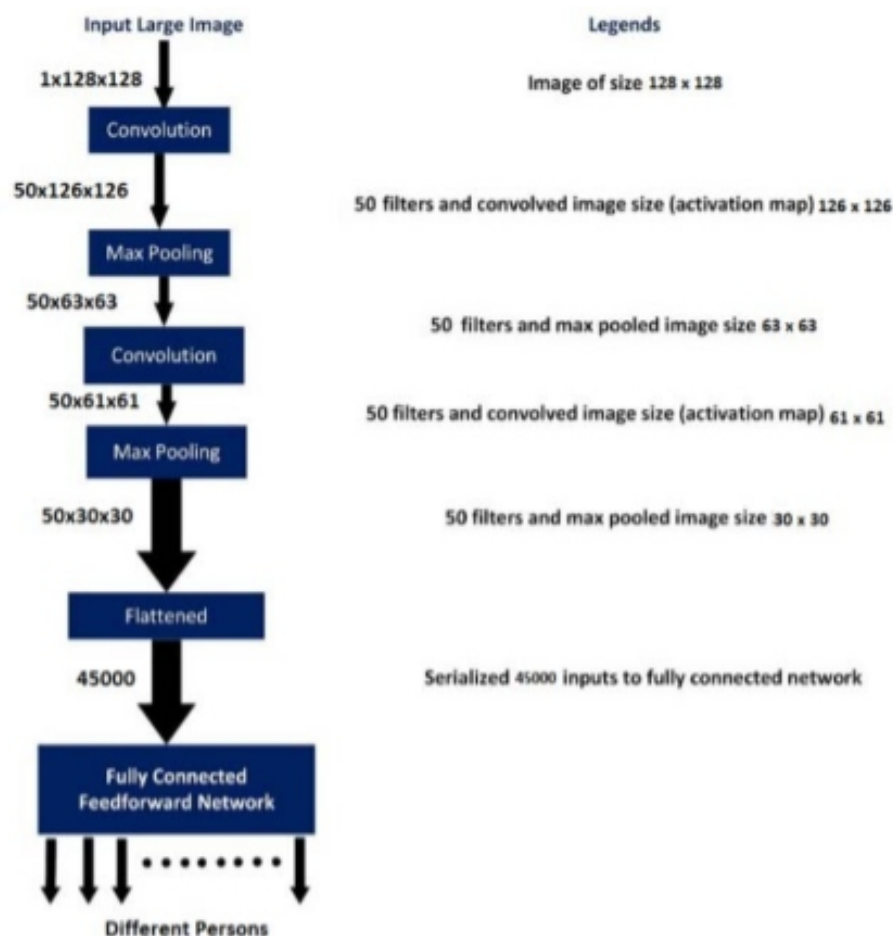2. Then a convolution layer with 50 kernels each of size 3*3 (this is also the size of the local receptive field) follows, which results in 50*126*126 hidden feature neurons.
3. The next step is to perform max pooling with a max pooling kernel size 2*2 with a stride of 2 (no overlapping) across all the 50 feature or activation maps and the result is a layer of 50*63*63 hidden feature neurons.
4. The process of convolution and max pooling is once again repeated with 50 kernels of size 3*3. With this, convolution layer yields a layer of 50*61*61 hidden feature neurons, while the max pooling layer yields a layer of 50*30*30 hidden feature neurons.
5. This is now flattened to a size of 45000 pixels at the next layer.
6. The final layer is a fully connected layer. So this layer connects every neuron from the flattened max pooled layer to every one of the 10 output neurons for the initial stage of learning and recognitionfor each CNN model (as there are 10 classes for each model in the first case). For the second stage of CNN learning and recognition, the number of output neurons is 25 in case of the face CNN, 40 in case of the fingerprint CNN and 36 in case of the iris CNN. [5]

Classifier Performance Evaluation

Holdout Method

In order to obtain a good measure of the performance of a classifier, it is necessary that the test dataset has approximately the same class distribution as the training dataset. The training dataset is that portion of the available labelled examples that is used to build the classifier. The test dataset is that portion of the available labelled examples that is used to test the performance of the classifier. In holdout method, the reason for keeping aside some of the labelled examples for a test dataset is to test the performance of the constructed classifier. [4, 7, 9, 11]

Confusion Matrix

A confusion matrix is a table that describes the performance of a classifier model on a set of test data for which the actual or target labels are known. This allows the visualization of the performance of the

classifier model. A good classifier can be detected by observing the values along the principle diagonal of the confusion matrix i.e. this principle diagonal must contain high values as this will prove that most of the testing data objects have been correctly categorized into their appropriate classes. However, for quantitative analysis, the accuracy of the classifier can be measured in terms of other metrics that are derived from the confusion matrix. [4]

		Actual Class	
		Class X	Class Y
Predicted Class	Class X	a	b
	Class Y	c	d

Figure 7. Confusion Matrix (2 classes)

In the figure of the confusion matrix shown above, there are only two classes - Class X and Class Y. The terms a, b, c and d are defined as follows:

- 'a' denotes the number of inputs that actually belong to class X and are also correctly predicted to belong to Class X.
- 'b' denotes the number of inputs that actually belong to class Y but are wrongly predicted to belong to Class X.
- 'c' denotes the number of inputs that actually belong to class X but are incorrectly predicted to belong to Class Y.
- 'd' denotes the number of inputs that actually belong to class Y and are also correctly predicted to belong to Class Y.

Most performance metrics are computed from the confusion matrix.

Accuracy

The accuracy of a classifier is the probability of its correctly classifying records in the test dataset. In practice, accuracy is measured as the percentage of records in the test dataset that are correctly classified by a classifier. If there are only two classes (X and Y), then the accuracy is computed as:

$$\text{Accuracy} = (a+d) / (a+b+c+d)$$

as shown in the confusion matrix above. [4, 9, 16]

Precision

The precision of a classifier is the probability of records actually being in class Positives if they are classified as being in class Positives. [16]

$$\text{Precision} = a / (a+b)$$

Recall

The recall of a classifier is the probability that a record is classified as being in class Positives if it is actually belongs to class Positives. [16]

$$\text{Recall} = a / (a+c)$$

F-Score

F-Score is the harmonic mean of precision and recall. [16]

Performance evaluation of the classifier

For the performance evaluation of the classifier, when we used Holdout method we were able to test patterns which were not included in training pattern set. In evaluation of a classifier with accuracy metric the overall performance of the classifier is reflected irrespective of the individual performance evaluation for each and every class or category.

IV. EXPERIMENTAL RESULTS

The present section tabulates the data obtained as results of the conducted experiments.

1. PLATFORM

The experiments are performed in a system having a dual-core processor of type Intel® Core™ i3-7020U CPU with a clock rate of 2.3 GHz and equipped with 4GB RAM and 1TB of Hard Disk space, that runs on Microsoft Windows 10 operating system. Initially the three CNN's are trained and tested using only 10 subjects or persons with 40 epochs for training on this platform. The confusion matrices are derived by considering this initial case of 10 subjects. Later on, the CNN models are trained using larger datasets and with more epochs in order to increase the accuracy of the CNN classifiers. This second stage of CNN learning and recognition has been carried out on Google Colaboratory ("Hello Colaboratory", https://colab.research.google.com) that provides free GPU and TPU for a maximum time of 12 hours per session via online cloud services. The GPU that can be accessed via this platform is an Intel® Xeon® GPU model that has 13 GB RAM, 33 GB memory and clock rate of 2.2 GHz.

2. BENCHMARKED DATABASE

A. Dataset for Face CNN

A part of the famous benchmarked databases called "Labelled Faces in the Wild" (LFW) is used. There are 5749 subjects with total number of 13,244 images. Out of these, 10 subjects are chosen for the initial

stage of CNN learning and recognition. Later on, the CNN model is trained using 25 people from the original dataset. The link to the database is: https://www.kaggle.com/jessicali9530/lfw-dataset

B. Dataset for Fingerprint CNN

Fingerprint images from FVC2002 database are used. There are 4 databases in this set. In each of the four databases, there are 10 subjects. The link to the FVC2002 fingerprint database is: http://bias.csr .unibo.it/fvc2002/databases.asp

C. Dataset for Iris CNN

The CUHK Iris image dataset obtained from the archives of the Chinese University of Hong Kong has been used here. There are 36 subjects. The link to the database is: http://www.mae.cuhk.edu.hk/~cvl /main_database.ht m

3. PERFORMANCE EVALUATION RESULTS

Performance of the CNN Classifier for Facial Recognition

Performance of the facial CNN classifier in terms of the accuracy, precision, recall and F-score has been evaluated with the help of confusion matrix. For this initial stage of training, the Face CNN model has been trained using only 10 subjects each having a total variation of 10 images and has been tested on the same 10 subjects using 5 of the images for each test subject.

Confusion Matrix

		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
	P1	3	0	1	0	0	0	1	0	0	0
	P2	0	2	0	1	0	0	1	0	1	0
	P3	0	0	0	0	1	1	0	1	1	0
	P4	1	0	0	3	0	0	0	0	0	1
	P5	0	1	0	0	4	0	0	0	0	0
	P6	1	0	0	1	0	2	0	0	1	0
	P7	0	1	1	0	1	0	1	0	0	1
	P8	0	0	0	1	0	1	0	3	0	0
	P9	0	0	0	0	0	1	1	0	2	0
	P10	0	0	1	0	1	0	0	1	0	2

Actual Class →, Predicted Class ↓

Figure 8. Confusion Matrix of Face CNN

Performance Evaluation Table

Person	Precision	Recall	F-score
P1	0.60	0.60	0.60
P2	0.40	0.40	0.40
P3	0.20	0.25	0.22
P4	0.60	0.50	0.54
P5	0.80	0.57	0.66
P6	0.40	0.40	0.40
P7	0.20	0.25	0.22
P8	0.60	0.60	0.60
P9	0.40	0.40	0.40
P10	0.40	0.50	0.44

Figure 9. Performance Metrics of Face CNN

Accuracy of Face Classifier = 41.60%

Finally, the Face CNN classifier is trained with 25 different subjects with each having 10 samples for a time of 60 epochs. It is then tested using the test samples of same 25 persons. It is noted that the accuracy of the classifier increases with increase in the dataset as well as the number of epochs.

The observed accuracy of the classifier = 85.40%.

Performance of the CNN Classifier for Fingerprint Recognition

Performance of the fingerprint CNN classifier in terms of the accuracy, precision, recall and F-score has been evaluated with the help of confusion matrix. For this initial stage of training, the Fingerprint CNN model has been trained using only 10 subjects each having a total variation of 8 images and has been tested on the same 10 subjects using 4 of the images for each test subject. Confusion Matrix

Predicted Class \ Actual Class	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
P1	2	0	1	0	0	0	1	0	0	0
P2	0	3	0	1	0	0	0	0	0	0
P3	0	0	2	0	0	1	0	0	1	0
P4	1	0	0	3	0	0	0	0	0	0
P5	0	0	0	0	4	0	0	0	0	0
P6	1	0	0	1	0	1	0	0	1	0
P7	0	0	1	0	0	0	3	0	0	0
P8	0	0	0	0	0	1	0	3	0	0
P9	0	1	0	1	0	0	0	0	2	0
P10	0	0	0	0	1	0	0	1	0	2

Figure 10. Confusion Matrix of Fingerprint CNN

Performance Evaluation Table

Person	Precision	Recall	F-score
P1	0.50	0.50	0.50
P2	0.75	0.75	0.75
P3	0.50	0.50	0.50
P4	0.75	0.50	0.60
P5	1.00	0.80	0.88
P6	0.25	0.33	0.28
P7	0.75	0.75	0.75
P8	0.75	0.75	0.75
P9	0.50	0.50	0.50
P10	0.50	1.00	0.66

Figure 11. Performance of Fingerprint CNN

Accuracy of the Fingerprint Classifier = 62.5% Finally, the Fingerprint CNN classifier is trained with 40 different subjects with each having 8 samples for a time of 60 epochs. It is then tested using the test samples of same 40 persons. It is noted that the accuracy of the classifier increases with increase in the dataset as well as the number of epochs.

The observed accuracy of the classifier = 75.36%.

Performance of the CNN Classifier for Iris Recognition

Performance of the iris CNN classifier in terms of the accuracy, precision, recall and F-score has been evaluated with the help of confusion matrix. For this initial stage of training, the Iris CNN model has been trained using only 10 subjects each having a total variation of 8 images and has been tested on the same 10 subjects using 4 of the images for each test subject.

Confusion Matrix

		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
	P1	3	0	0	0	0	0	1	0	0	0
	P2	0	1	0	1	0	0	0	0	1	1
	P3	0	0	2	0	0	1	0	0	1	0
	P4	0	0	0	4	0	0	0	0	0	0
	P5	0	0	0	0	3	0	0	1	0	0
	P6	1	0	0	0	0	2	0	0	1	0
	P7	0	0	1	1	0	0	2	0	0	0
	P8	0	0	0	0	0	1	0	3	0	0
	P9	0	1	0	1	0	0	0	0	2	0
	P10	0	0	0	0	1	0	0	1	0	2

Figure 12. Confusion Matrix of Iris CNN

Performance Evaluation Table

Person	Precision	Recall	F-score
P1	0.75	0.75	0.75
P2	0.25	0.50	0.33
P3	0.50	0.66	0.57
P4	1.00	0.57	0.73
P5	0.75	0.75	0.75
P6	0.50	0.50	0.50
P7	0.50	0.66	0.57
P8	0.75	0.60	0.66
P9	0.50	0.40	0.44
P10	0.50	0.66	0.57

Figure 13.Performance Metrics of Iris CNN

Accuracy of the Iris Classifier = 60%

Finally, the Iris CNN classifier is trained with 36 different subjects with each having 7 samples for a time of 60 epochs. It is then tested using the test samples of same 36 persons. It is noted that the accuracy of the classifier increases with increase in the dataset as well as the number of epochs.

The observed accuracy of the classifier = 82.14%.

4. LEARNING AND PERFORMANCE EVALUATION TIME

Learning and Performance Evaluation time for Facial CNN

There were 10 subjects i.e. 10 different persons. The performance evaluation of the system was conducted with the same 10 different subjects. They are tabularized below:

Performance time	Evaluation time	Total time
13.19 minutes	7.12 minutes	20.31 minutes

In the latter stage, the Face CNN classifier is trained with 25 different subjects with each having 10 samples for a time of 60 epochs. It is then tested using the test samples of same 25 persons. The dataset has been partitioned for training and testing as per the hold-out method. Thereafter the time for training the CNN with 250 images and the evaluation time for recognizing test images of 25 subjects are noted:

Performance Time	Evaluation Time	Total Time
8.67 minutes	5.15 minutes	13.82 minutes

Learning and Performance Evaluation Time for Fingerprint CNN

There were 10 subjects i.e. 10 different persons. Thereafter the performance evaluation of the system was conducted with the same 10 different subjects. They are tabularized below:

Performance Time	Evaluation Time	Total Time
10.32 minutes	5.27 minutes	15.59 minutes

Later on, the Fingerprint CNN classifier is trained with 40 different subjects with each having 8 samples for a time of 60 epochs. It is then tested using the test samples of same 40 persons. The dataset has been partitioned for training and validation as per the hold- out method. Thereafter the time for training the CNN with 288 images and the evaluation time for recognizing test images of 40 subjects are noted:

Performance Time	Evaluation Time	Total Time
19.04 minutes	11.52 minutes	30.56 minutes

Learning and Performance Evaluation Time for Iris CNN

There were 10 subjects i.e. 10 different persons. Thereafter the performance evaluation of the system was conducted with the same 10 different subjects or categories. They are tabularized below:

Performance Time	Evaluation Time	Total Time
11.22 minutes	4.09 minutes	15.31 minutes

Later on, the Iris CNN classifier is trained with 36 different subjects with each having 8 samples for a time of 60 epochs. It is then tested using the test samples of same 36 persons. The dataset has been partitioned for training and validation as per the hold- out method. Thereafter the time for training the CNN with 288 images and the evaluation time for recognizing the test images of 36 subjects are noted:

Performance Time	Evaluation Time	Total Time
17.31 minutes	7.03 minutes	24.33 minutes

V. CONCLUSION

In the present paper, we have designed and developed one biometric-based person identification system using CNN model classifier. The performance evaluation of the CNN based classifier in terms of confusion matrix and its associated metrics is satisfactory. The person recognition based on biometric authentication is effective, efficient and most promising for future uses. The work can be augmented by integrating the conclusions of the different classifiers with different biometrics such as fingerprints, faces, palm-prints, wrist-veins, iris, handwriting etc. to build up one super classifier This is the future scope of the work and a super classifier authentication would be far more accurate than ordinary single classifier classification.

REFERENCES

[1] François Chollet, "Deep Learning with Python", Manning Publications, 2017.

[2] Ian Goodfellow, "Deep Learning", MIT Press.

[3] Sarker G., "A Treatise on Artificial Intelligence", ISBN : 978-93-5321-793-8 ; 1, 325, 2018

[4] Arun K. Pujari, "Data Mining Techniques", Universities Press

[5] Sarker, G. (2018) Some Studies on Convolution Neural Network - International Journal of Comuter Applications (IJCA), Foundations of Computer Science, New York, Vol. 182 - No. 21, DOI 10.5120/ijca201891, pp 13-22, Oct. 2018.

[6] Sarker, G. Ghosh S. (2018) A Convolution Neural Network for Optical Character Recognition and Subsequent Machine Translation - International Journal of Comuter Applications (IJCA), Foundations of Computer Science, New York, Vol. 182 - No. 30, DOI 10.5120/ijca2018918203, pp 23-27, Dec. 2018.

[7] Kundu, S., Sarker, G.(2013),A Modified Radial Basis Function Network For Occluded Fingerprint Identification and Localization ,IJCITAE, Serial Publications,7(2), 103- 109.

[8] Sarker, G.,Bhakta, D.(2015),An Unsupervised OCA based RBFN for Clear and Occluded Face Identification, Intelligent Computing and Applications, Advances in Intelligent Systems and Computing 343,DOI 10.1007/978- 81- 322-2268-2_3.19-25.

[9] Kundu, S., Sarker, G. (2015), A Modified SOM Based RBFN for Rotation Invariant Clear and Occluded Fingerprint Recognition, Intelligent Computing and Applications, Advances in Intelligent Systems and Computing 343,DOI 10.1007/978-81-322-2268-2_2.

[10] Kundu, S., Sarker, G. (2016), A Multi-level Integrator with Programming Based Boosting for Person Authentication using Different Biometrics, Journal of Information Processing Systems, Volume 14, Issue 5, pp 1114- 1135 (SCOPUS & ESCI indexed).

[11] Sarker, G. Kundu, S., (2013),A Modified Radial Basis Function Network for Fingerprint Identification and Localization ,International Conference on Advanced Engineering and Technology held at Kolkata on 15th Feb. 2013, pp. 26-31.

[12] Bhakta, D., Sarker, G.(2013),A Radial Basis Function Network for Face Identification and Subsequent Localization International Conference on Computer Science & Information Technology (ICCSIT), pp.1-6, 10th March 2013, ISBN 978-93-82208-70-9.

[13] K. Roy, Sarker G. (2013),A Location Invariant Face Identification and Localization with Modified RBF Network , International Conference on Computer and Systems ICCS-2013, 21-22 September, 2013, pp − 23-28, Bardhaman.

[14] Bhakta, D., Sarker, G.(2013),A Rotation and Location Invariant Face Identification and Localization with or without Occlusion using modified RBFN , Proceedings of the 2013 IEEE Second International Conference on Image Information Processing (ICIIP − 2013), DOI: 10.1109/ICIIP.2013.6707649, pp − 533-538.

[15] Kundu, S., Sarker, G.(2014),A Modified RBFN Based on Heuristic Based Clustering for Location Invariant Fingerprint Recognition and Localization with and without Occlusion ,International Conference for Convergence of Technology – 2014, DOI: 10.1109/I2CT.2014.7092281, IEEE, pp 1-6.

[16] Kundu, S., Sarker, G.(2014),A Modified BP Network using Malsburg Learning for Rotation and Location Invariant Fingerprint Recognition and Localization with and without Occlusion ,2014 Seventh International Conference on Contemporary Computing (IC3), DOI: 10.1109/IC3.2014.6897244, 617-623.

[17] Kundu S., Sarker G. A Person Identification System with Biometrics using Modified RBFN based Multiple Classifier – 2016 International Conference on Control, Instrumentation, Energy & Communication (CIEC), held at Kolkata, January 2016.

[18] Huiying Hu ; Syed Afaq Ali Shah ; Mohammed Bennamoun ; Michael Molton, "2D and 3D face recognition using convolutional neural network", TENCON – IEEE Region 10 Conference, November 2017.

[19] Hurieh Khalajzadeh, Mohammad Mansouri and Mohammad Teshnehlab, "Face Recognition using Convolutional Neural Network and Simple Logistic Classifier", Online Conference on Soft Computing and Industrial Applications Anywhere on Earth, December 2012.

[20] Zhonghua Liu, Zhong Jin, Zhihui Lai, Chuanbo Huan, Minghua Wan, "Face Recognition Based on Wavelet Transform, Singular Value Decomposition and Kernel Principal Component Analysis", Chinese Conference on Pattern Recognition, 2008.

Instructions for Authors

Essentials for Publishing in this Journal

1   Submitted articles should not have been previously published or be currently under consideration for publication elsewhere.

2 Conference papers may only be submitted if the paper has been completely re-written (taken to mean more than 50%) and the author has cleared any necessary permission with the copyright owner if it has been previously copyrighted.

3 All our articles are refereed through a double-blind process.

4 All authors must declare they have read and agreed to the content of the submitted article and must sign a declaration correspond to the originality of the article.

Submission Process

All articles for this journal must be submitted using our online submissions system. http://enrichedpub.com/ . Please use the Submit Your Article link in the Author Service area.

———————————————————————

Manuscript Guidelines

The instructions to authors about the article preparation for publication in the Manuscripts are submitted online, through the e-Ur (Electronic editing) system, developed by **Enriched Publications Pvt. Ltd**. The article should contain the abstract with keywords, introduction, body, conclusion, references and the summary in English language (without heading and subheading enumeration). The article length should not exceed 16 pages of A4 paper format.

Title

The title should be informative. It is in both Journal's and author's best interest to use terms suitable. For indexing and word search. If there are no such terms in the title, the author is strongly advised to add a subtitle. The title should be given in English as well. The titles precede the abstract and the summary in an appropriate language.

Letterhead Title

The letterhead title is given at a top of each page for easier identification of article copies in an Electronic form in particular. It contains the author's surname and first name initial .article title, journal title and collation (year, volume, and issue, first and last page). The journal and article titles can be given in a shortened form.

Author's Name

Full name(s) of author(s) should be used. It is advisable to give the middle initial. Names are given in their original form.

Contact Details

The postal address or the e-mail address of the author (usually of the first one if there are more Authors) is given in the footnote at the bottom of the first page.

Type of Articles

Classification of articles is a duty of the editorial staff and is of special importance. Referees and the members of the editorial staff, or section editors, can propose a category, but the editor-in-chief has the sole responsibility for their classification. Journal articles are classified as follows:

Scientific articles:

1. Original scientific paper (giving the previously unpublished results of the author's own research based on management methods).

2. Survey paper (giving an original, detailed and critical view of a research problem or an area to which the author has made a contribution visible through his self-citation);

3. Short or preliminary communication (original management paper of full format but of a smaller extent or of a preliminary character);

4. Scientific critique or forum (discussion on a particular scientific topic, based exclusively on management argumentation) and commentaries. Exceptionally, in particular areas, a scientific paper in the Journal can be in a form of a monograph or a critical edition of scientific data (historical, archival, lexicographic, bibliographic, data survey, etc.) which were unknown or hardly accessible for scientific research.

Professional articles:

1. Professional paper (contribution offering experience useful for improvement of professional practice but not necessarily based on scientific methods);

2. Informative contribution (editorial, commentary, etc.);

3. Review (of a book, software, case study, scientific event, etc.)

Language

The article should be in English. The grammar and style of the article should be of good quality. The systematized text should be without abbreviations (except standard ones). All measurements must be in SI units. The sequence of formulae is denoted in Arabic numerals in parentheses on the right-hand side.

Abstract and Summary

An abstract is a concise informative presentation of the article content for fast and accurate Evaluation of its relevance. It is both in the Editorial Office's and the author's best interest for an abstract to contain terms often used for indexing and article search. The abstract describes the purpose of the study and the methods, outlines the findings and state the conclusions. A 100- to 250-Word abstract should be placed between the title and the keywords with the body text to follow. Besides an abstract are advised to have a summary in English, at the end of the article, after the Reference list. The summary should be structured and long up to 1/10 of the article length (it is more extensive than the abstract).

Keywords

Keywords are terms or phrases showing adequately the article content for indexing and search purposes. They should be allocated heaving in mind widely accepted international sources (index, dictionary or thesaurus), such as the Web of Science keyword list for science in general. The higher their usage frequency is the better. Up to 10 keywords immediately follow the abstract and the summary, in respective languages.

Acknowledgements

The name and the number of the project or programmed within which the article was realized is given in a separate note at the bottom of the first page together with the name of the institution which financially supported the project or programmed.

Tables and Illustrations

All the captions should be in the original language as well as in English, together with the texts in illustrations if possible. Tables are typed in the same style as the text and are denoted by numerals at the top. Photographs and drawings, placed appropriately in the text, should be clear, precise and suitable for reproduction. Drawings should be created in Word or Corel.

Citation in the Text

Citation in the text must be uniform. When citing references in the text, use the reference number set in square brackets from the Reference list at the end of the article.

Footnotes

Footnotes are given at the bottom of the page with the text they refer to. They can contain less relevant details, additional explanations or used sources (e.g. scientific material, manuals). They cannot replace the cited literature.

The article should be accompanied with a cover letter with the information about the author(s): surname, middle initial, first name, and citizen personal number, rank, title, e-mail address, and affiliation address, home address including municipality, phone number in the office and at home (or a mobile phone number). The cover letter should state the type of the article and tell which illustrations are original and which are not.

Address of the Editorial Office:

Enriched Publications Pvt. Ltd.
S-9,IInd FLOOR, MLU POCKET,
MANISH ABHINAV PLAZA-II, ABOVE FEDERAL BANK,
PLOT NO-5, SECTOR -5, DWARKA, NEW DELHI, INDIA-110075,
PHONE: - + (91)-(11)-45525005

Note