

“© 2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

Application of Data Science for Controlling Energy Crises: A Case Study of Pakistan

Saif Ullah⁽¹⁾, Muhammad Asif^(1*), Shahbaz Ahmad⁽¹⁾, Ulfat Imdad⁽¹⁾, Osama Sohaib⁽²⁾

1. Department of Computer Science, National Textile University Faisalabad.

E-mail: saif.pro.ds@gmail.com, asif@ntu.edu.pk, shahbaz@ntu.edu.pk, ulfatimdad@gmail.com

2. School of Systems, Management and Leadership, Faculty of Engineering and Information Technology University of Technology Sydney, Australia

*corresponding author

[e-mail: Osama.Sohaib@uts.edu.au]

Abstract—Today Pakistan is facing numerous challenges for the interconnection of local energy resources and for balanced energy policies. Data Science, Big Data, Artificial Intelligence (AI), IoT and Cloud computing draws our focus towards controlling energy crises in terms of smart energy generation, consumption and to overcome causes of energy crises. To make a conclusion valuable we have to extract significant value from a large amount of data that's why data management plays a significant role. This Paper presents a review of energy sectors, energy resources, energy crises in Pakistan. It also presents the possible solution of energy crises with the help of data science application and the involvement of Big Data, Cloud computing, IoT and AI

Keywords— Energy Crisis, Energy Resources, Energy Sector, Data Science, Big Data, Cloud Computing, Artificial Intelligence

I. INTRODUCTION

In today's world, data science plays a vital role in different fields of technology and shows an astonishing result, advancement in knowledge allows us to improve or automate the infrastructure of life with technology. Over the former era, the power grid actions have turned into more stressed due to rising customer request and less protected by the combination of irregular renewable assets. Energy is the salvation of a homeland. Energy in the shape of power is vital for the financial growth of any homeland and now mostly known as a significant product designed for homo sapiens. Energy is the basic necessity of economic engine, as well as the wheels of business, food production, and industry. As a collective scenario, utilization of energy per consumer is a crucial sign of the value of life cycle of the nations. sadly, in Pakistan, in spite of big demands and hyperbole by all governments, actual solutions to encounter the energy supplies for the homeland have never been accomplished. As a result, Pakistan's commercial, business, and public development have been prominently reserved. Energy ingestion is a key to business economy and richness of the society in a homeland because energy is the main influence for almost wholly social actions and improvements. Due to growing universal residents and grasping routines of the society, the energy assets are reducing quickly. Improvements in communication, information technologies (IT), regulator, and related applications to energy manufacturing have provided the utensils for revolutionizing

the old-fashioned energy grid. The development of the Smart Grid represents a more, flexible, circulated, and collaborative role for the final-user in the time-to-time operations of the organization [1]. Pakistan is amongst those emerging countries which are facing stark deficiencies of energy nowadays. The demand-supply hole of Pakistan's energy sector is growing mostly due to the extraordinary growth rate of the population which is presently projected about 2%, the demand is growing while resource has shown development at snail's pace [2]. The sluggish progress of energy resource is due to deprived infrastructure, giant capital involved to shape supply assets and postponed policies on the application level. Pakistan has generous energy assets but massive regions of the country are destitute of commercial energy access, in future, there is a dire necessity to improve all forms of accessible resources to moderate demand-supply hole [3].

II. ENERGY IN PAKISTAN

A. Energy Sources

The energy sectors of Pakistan are mostly beneath explicit and ambiguous control by both sectors government and private. Pakistan's energy framework is embryonic, inadequate and badly managed. Domestic, Transport, Industrial, Agriculture, Commercial, and Government consumption sectors are some of the main energy sectors of Pakistan.

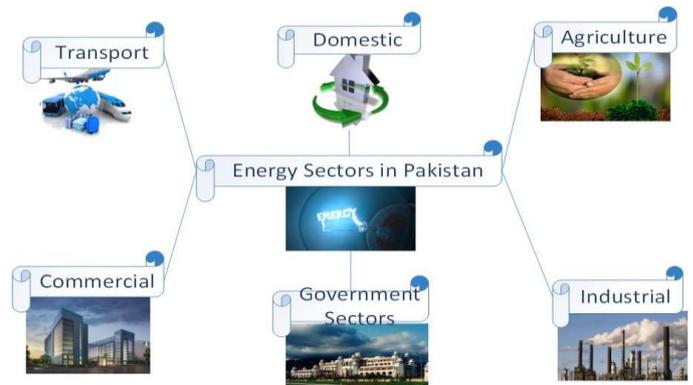


Figure 1 Energy Sectors of Pakistan

B. Energy Sources

Oil, Gas, LPG, Coal, Hydroelectricity, Nuclear Electricity are the major energy foundations in Pakistan [4].



Figure 2 Energy Sources in Pakistan

C. Energy Sector Organizations

The following organizations are playing an important role in helping energy transmission, generation, and circulation in Pakistan: [5]

1) Organization of Water and Power

The managerial support of Pakistan Government is the federal minister of Water and Energy, for wholly problems regarding electricity production, broadcast, and circulation, regulation, valuing and ingestion inside the homeland. Organization of water and energy executes these tasks with the help of its numerous line workstations and appropriate independent groups [5].

2) National Electric Power Regulatory Authority

The National Electric Power Regulatory Authority (NEPRA) was recognized by an act of the Legislature. NEPRA was designed as an autonomous supervisor, in demand to authorize a commercially-oriented, modest and crystal-clear energy sector in Pakistan.

The key tasks of this specialist Authority contain, production licenses allotting, distribution and transmission of power; approving power achievement and resource programmers of the utility corporations; launching and executing principles to ensure protection, excellence, and appropriate secretarial of process and supply of power to clients; and defining fares for transmission, bulk production and retail distribution of electric power [5].

3) Private Power Infrastructure Board

Th PPIB performs as a 'single' initiator for conservative private sector energy production schemes. This includes schemes containing RE Hydel scheme having a capacity of greater than 50MW volume in our homeland. Private Power Infrastructure Board represents Azad Jammu Kashmir and provinces of Pakistan.

4) Provincial and AJK agencies

AJK (Azad Jammu Kashmir) and Provincial governments encourage besides backing the execution of sustainable resources schemes within the physical authority. These tasks are completed on particular bases and sometimes with the help of alliance with AEDB. For these tasks altered approaches are engaged, like as by authorizing besides generating responsiveness of reprocessing, advancing and helping in the distribution of terrestrial use human privileges (e.g., aimed at airstream ranches), besides eliminating further disorders. As a result, this may delay advancement in their progress. In Azad Jammu Kashmir (AJK) and each province. [5]

5) Power utilities

In Pakistan, the energy distribution facility was commenced by altered organizations. Public and Private segments have organizations, in different sections. In 1958 with the help of WAPDA Act Water and Power Development Authority (WAPDA) was established, which delivers the joined besides the synchronized improvement of energy assets and water assets [5].

D. Causes of Energy Crisis in Pakistan

The major causes of energy crises in Pakistan are: [6]

1) Failure to enhance the producing capability

The elementary reason is the absolute failure of the earlier administration to raise the supply of energy to retain speed with the rising demand. While the mounted producing capability had improved 53 percent between 1994 and 1999 (from 11,320 MW to 17,400MW), it improved by simply 12 percent between 1999 and 2007, near 19,420 MW. Later 2007 to 2015 no main development was accomplished and assisted in Pakistan to the overwhelmed energy crisis. Even that negligible growth was due to the accomplishment of Ghazi Barotha Project started in the 1990s, earlier Gen Musharraf took over [6].

2) Problem of circular debt

Circular debt discusses to the debt dominant between altered sectors of the government over the power consumption charges, for example, the organizations which supply power to all the government sectors and also to KESC and PEPCO but they all do not pay for the electricity they use.

One of the key causes for the severe shortfall in the production of thermal power is the problem of the "circular debt" which the present government inherited from the earlier government. In 2007, the government did not pay the powerhouses for the grant that was being provided to customers. In 2014 government free few volumes of the circular debt due to force straddling from altered areas of life. In 2015 circular debt touched Rs.600 billion [6].

3) Losses in transmission and distribution

Very dense line losses in distribution and transmission because of ancient and ailing retained transmission and distribution system, projected at over 20 percent compared to 8 to 10 percent in further countries. Enormous scale stealing of energy as obviously exposed by the rising variance between units produced or acquired and those rewarded for [6].

4) *Wastage of Energy*

Industry waste most of the energy. Industry consumes almost 30 percent of total energy due to less organized structures and other practices. For example, Chinese textile mills consume 30 percent less electricity because they use water moderately controlled by solar panels in their tanks. Misuse of energy by the transportation sector (consuming 28 percent of overall energy) due to ancient and ailing tuned engines [6].

5) *Domestic and household consumption*

A large amount of consumption of total electricity is Domestic and household consumption which uses 45 percent also describes inefficient and excessive uses of air conditioners, lights, and large-scale illumination on different occurrences. The difficulties defined above expose many operational faults in our energy system. These contain over-reliance on imported energy, insufficient political determination, restricted financial backing and very pathetic implementation ability.

6) *Aging of the equipment*

Significant aim credited to the power deficiency is due to getting old of the producing machinery that might not produce power according to the strategy obligation. We honestly consider a severe attention could be specified designed for over-all renovation besides conservation of current machinery to retain all machinery in a decent operational situation.

7) *Tremendous Fuel Prices*

Crude oil prices have varied quickly. As a result, the production in which thermal units plays a vital role is estimated excessive costs. KESC besides WAPDA as soon as obtaining energy at greater prices is not enthusiastic to face loss by keeping an electricity on retailing. Consequently, they do not transfer towards the universal objection of power shortfall. One humble answer is that they raise the power prices.

8) *No coal extraction*

Pakistan is consecrated with a great volume of the coal. No thoughtful effort is ready to realize the coal as a key item for energy production. This is grumbled that worth of the coal is lower. Conversely, convenient answers are accessible in a mean to burn different coal categories. The administration is looking towards the private segment to contribute to this

scenario. According to our opinion, Administration himself must move at frontward and contribute to the installment of power plants at the places of coal mines.

Baloch MH and his copartners summarize the energy condition of the country with the features of wind energy potential in the southern section of the country in [30]. Siddique [31] explains and studied the progress besides the progress in the wind energy sector to our country and also gives some beneficial strategies that support us in power generation and power mix. Data of wind speed for different provinces is studied in [32] with the capability of assimilating grid with wind energy.

E. *Effects of Energy Crisis in Pakistan*

Most of the industrial units do not have their private generating capacity. In detail, most of the initiatives cannot work economically when electricity is accessible for merely rare hours in a day

Thousands of labors have been positioned off in the face of extraordinary food expenses, extraordinary increase and rising joblessness. This will also have a very undesirable influence on the country's export which is desired dreadfully to overwhelmed the economic disaster.

Textile business in Pakistan is going through one of the roughest periods in eras. The universal collapse which smashes the universal textile really tough is not the single reason for worry. Thoughtful inner concerns containing energy disaster affected Pakistan's textile business very desperately. The extraordinary expense of production causing an immediate increase in the energy cost has been the crucial reason for anxiety for the business Pakistan's textile industry is deficient in research & advancement. The production proficiency is very squat due to superseded machinery and technology.

III. INTRODUCTION TO DATA SCIENCE

During several centuries, high-tech utensils take advanced in a varied variety of areas, Power Proficiency besides Administration stays no exemption. Improvements inside numerous zones of new technology and Technology (ICT) Infrastructures, for instance, Robotics as well as Computerization, Real-time Supervision, Smart Metering as well as Data Science, give a marvelous impression regarding this arena. As long as it is well recognized, Data Science constructs outlines as well as procedures in a meanwhile to determine information, generate useful insights and detect patterns and forecasts of wide-scale information. Data Science incorporates entire deep analysis progression, that initiates due to data abstraction besides munging as well as prolongs towards data deep analysis, explanation besides definition. Affecting outcomes are the forecasting regarding innovative principles besides their perception. Data Science is the combination of mathematical analysis with different tools of IT (Information technology) and many others.

Conversely, originating perceptions with data is no longer attained by means of similar procedures. Professional essentially organize and understand the data in a meanwhile to gain appreciated information. As shown in Fig-3, this procedure initiates thru the gathering of natural or unprocessed data. Afterwards, cleaning the data is an essential part, besides the selection of the subdivision which contains significant data. For this reason, Professional implement strainers on that data and articulates queries with the help of which we exclude extraneous info. At present phase, probably possible that extra causes of data may be assimilated besides bonded on that inventive info or data in a meanwhile to deliver additional information. When our data is designed for practice, an investigative analysis can support to choose that whatever approaches and algorithms used are supreme operative in a meanwhile to attain the wanted information. The concluding procedure will take us towards the firm of penalties with the help of which decision-making can be a monitor, that may be dependent on conception.

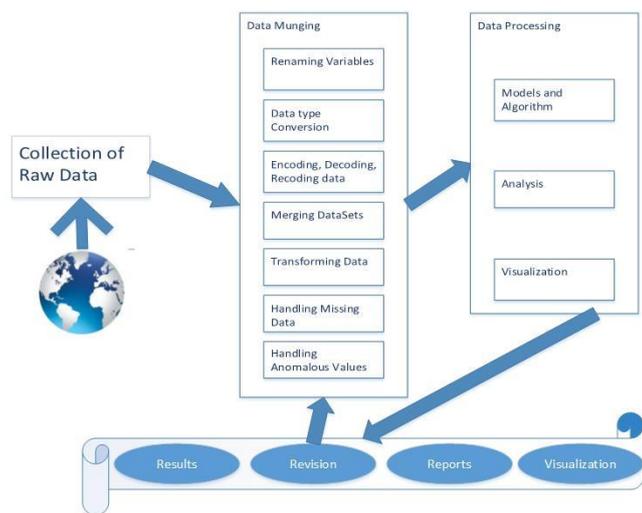


Figure 3 Data Science Process

For more practical perception, Data Science includes different combinations of algo besides utensils that help in chasing altered objectives and advance at altered conditions. A few Supreme methodologies are a regression, clustering, classification besides affiliated rules digging. All these procedures or algorithms have been habitually implemented in Power or Energy management and optimization, some other algorithms are also beneficial in a way to providing classification and other techniques in energy management and optimization like sequence analysis and anomaly detection [7].

A. Introduction to Regression Algorithm

The basic formula of predictive model algo is regression analysis that explores an affiliation among an independent (objective) and the dependent variable (forecaster) this method is beneficial for forecasting, interval sequence modeling and verdict the contributing influence affiliation among the variables.

There is numerous assistance in using regression analysis. Regression Analysis specifies momentous relationships amongst independent variable and dependent variable. Regression Analysis identifies the stability of impression of numerous independent on a dependent variable [8].

B. Introduction to Linear Regression Algorithm

One of the hottest topics that people pick while learning predictive modeling is Linear Regression. Linear regression consists of independent variables and dependent variables. Independent variable can be continuous or discrete while dependent variables are continuous and the nature of the regression line is linear. Multiple regression grieves from autocorrelation (relation between elements), heteroscedasticity (having diverse variances), multicollinearity (predictor variables are themselves highly correlated). Linear Regression is very penetrating to Outliers (observation point that lies abnormal difference). Multicollinearity can escalation the difference between the constant approximations and helps in making the approximations more penetrating to petty fluctuations in the model. The main consequence is the insecurity of the coefficient evaluation. In the different situation of various independent variables, forwarding selection can help in driving, step wise approach, regressive exclusion and for a variety of Supreme momentous independent variables [8].

C. Introduction to Logistic Regression Algorithm

Logistic Regression is a probability to find the success rate of the event and to find failure rate of an event. We should use logistic regression when the nature of the dependent variable is binary in the form of 0 or 1 (True or False). In logistic regression, the value of Y varies from 0 to 1. Logistic Regression is extensively used for classification complications. A Linear relationship between dependent and independent variables doesn't require the Logistic Regression. In a way to predict odd ratio, it can apply a nonlinear log transformation that's why it can handle various types of relationships. To evade over right and below right, we should comprise all-important variables.

D. Decision Tree

A decision tree is an arrangement, where each inner (non-leaf) node indicates a test on an element, each division indicates the consequence of a test, an individual leaf (or terminal) node grasps a class label. The root node is the topmost node in a tree. There are numerous precise decision-tree algorithms.

E. Clustering

Partitioning of items into clusters that mainly based on their resemblance grade is known as clustering. [9]. Clustering is mainly unsupervised since there is no preceding information of different modules to which the entities may be allocated. Different modules of cluster analysis mainly depend upon the principles used to measure resemblance [10]. Connectivity model, centroid model, distribution model, density model and graph-based model are some important models. By using

different models or algorithm's we enhance our chosen principle.

IV. DATA SCIENCE APPLICATION FOR ENERGY CRISIS

These are possessions that nearly corporations do currently with data, which, if finished intelligently, are definitely well-intentioned of actuality incorporated beneath the classification of application of 'data science' for energy crises solutions

A. . Energy Demand Forecast

One of the burning issues both for utilities and energy users is the capability to forecast usage and expenses for apparent aims. Utilities would like to be talented to regulate supply vs consumption and energy buyers would like to regulate their costs. Expressive what's going to ensure is vital to that. Numerous merchandises these ages' try to use predictive analytics approaches to forecast energy utilization and highest demands. Humble applications use primarily averaging procedures. Advanced operations (ones that are praiseworthy of working beneath the data science classification) comprise software that trains above antique energy patterns, associate them with outside data such as functioning data, weather etc. and generate realistically precise forecasts of future demands. Sophisticated solutions contain even real-time enthusiastically fluctuating forecast, that aid ability executives and utilities make online verdicts. Fig shows the general philosophy of forecasting. With the help of this solution we can solve many of our problems but firstly we have to update our systems and introduce new technology to our old ancient systems. By applying this data science application in our energy system, we will soon overcome our energy crisis that affects our major sectors:

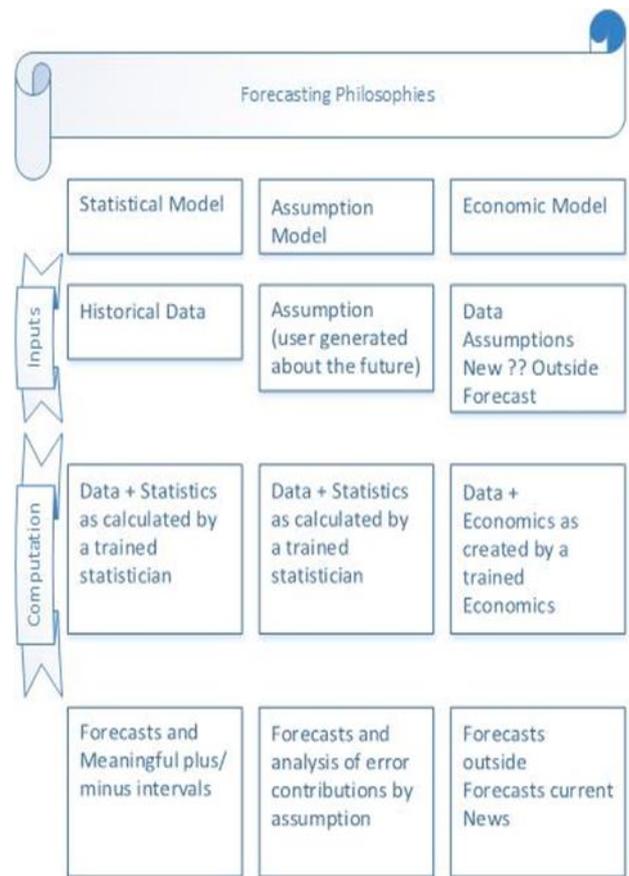


Figure 4 Forecasting Philosophies

B. Energy data disaggregation

Wastage of energy is the key factor of energy crises in Pakistan. Capability to fragmented your energy bill into energy convention is fundamental to recognize where energy waste occurs and increase energy efficiency. Disaggregation is an enactment of analytical approaches on topmost of accumulated energy data in demand to organize detailed usage designs. Smearing these techniques on a home owner's meter data can convey about how greatly energy is consumed by AC, how abundant is used for illuminations (lights) and exactly how much for refrigeration. Progressive operations can deliver this information in near real time or real time. Numerous utilities practice these performances to support their customers to save energy.

C. Real time device state classification

Solutions that manage energy at the device level is a comparatively new field. I.e. observer real-time energy at the specific device level, and deliver perceptions about energy or functioning proficiency out of that data. In this situation, advanced and progressive machine learning algorithms are used in real time to unconventionally classify functioning state of a device – is it off? Is it on? Is it idling? and further, out of its real-time energy ingesting data. Perceptive functioning state of each device can support identify irregularities (is device operational when it is invented too? Are these

machines working in sync?), that can be provided to facilitate and energy supervisors as sirens and support them to manage energy and functioning effectively in real time.

D. Building Operations

Nowadays, building management classifications produce an extensive amount of data. These data cover data regarding moisture, temperature, flow rate, energy, signals regulator, pressure, equipment grade, etc. Therefore, they may be evaluated and utilized to abstract operating instructions for finding building operations. In many situations, these rules are certainly decipherable IF-THEN rules, which can support to produce references for control approaches such as response appliances, and alterations of control procedures [11]. Such as May-Ostendorp et al. [12] abstract principles from disconnected prototypical predictive regulator sequence by applying different methods of classification. The authors utilize those instructions in order to accomplish near-optimal managing switch approaches for a building of mixed-mode throughout the chilling spell. Even though a range of techniques can be pragmatic to abstract IFTHEN instructions, classification technique can be used by main authors [13,14], predominantly decision tree algorithms [11,12]. In accumulation, classification models are operational utensils that are beneficial to forecast building operator relief below altered conservational circumstances [11]. Xiao et al. [15] ARs were used to identify the associations amongst the energy utilization of main mechanisms in an individual collection.

E. Importance of Smart Metering

Smart metering helps in the uninterrupted observation of power depletion with an observation to gaining an enhanced sympathetic of power generation, utilization besides transference phases. In the coming world, it will expressively advance the segment of decision-making, and similarly surely affect the performance of power of ending operators besides executives. It indicates a massive variation in the meters of gas and electricity subsequently it processes the accurate volume of expended energy also archives the precise period when depletion arises. Smart metering licenses the real-time consumed energy visualization. Users are capable of enhanced and recognize their particular energy performance. It brings safety beside confidentiality anxiety with them since users are indulged to contribute practice data [16]. Supplementary maintenance should be reserved so as not to classify specific users throughout the analysis manner. Data anonymization is the most regular solution for this delinquent [17].

V. BIG DATA

We live in a progressively restrained culture. The amalgamation of cloud computing, the Internet, and mobile devices is stirring continually superior servings of our lives and industries inside the data center. As a consequence, we are producing and must evaluate, massive and diverse flows of data. The performances established to do this procedure the form of expertise we discuss as "big data." This capability to find awareness outside ordinary human hold has wedged the imagination of the community. The particular description of

"big data" is foggy, although I, like numerous others, have thrashed to convey one:

"The data which surpasses the treating ability of conventional database such data is known as Bigdata. Big data is too immense, changes rapidly and doesn't even fit into the limits of our database design. In order to capture the important or valuable data from such data, we have to choose a nontraditional way to practice it"

Big Data has expanded the concentration of administrations and corporations. Such as, the Command of European and data of Europe's diligence (corporations just like as Orange, SAP and Siemens) their ambition is to reinforce the data region in a way to positioned Europe in the competition of the data at the forefront So they contracted to finance €2.5 billion in a partnership of public and private (PPP) [18]. Big data resolutions are booming and now a day's lead at the frontend in many businesses, just like Retail, Health, Education and energy management. [19]. The answers regarding energy are provided by companies with a modest benefit, and power production is no immunity. For illustration, Lesser reported as [20], 1000 petabytes of data are hypothetically produced when smart meters are fully systematized. Summarization and smarter analysis are needed for such an enormous flow of information, while Big data have a lot of things in this scenario. In statistic, IBM [21] precisely declares that the dares and occasions of Big Data as functional to energy optimization and management [10].

VI. ARTIFICIAL INTELLIGENCE AND IOT

The World recognizes an Internet of Things as a term in which not only people but also objects will be able to interact through the internet and are permanently connected. The IoT is estimated to adoptive an enormous amount of innovative technologies just like as health care, monitoring of the environment, and well-organized energy optimization and management in smart cities and smart homes. Rendering to Rifkin [22], the IoT can encourage more cooperative economy with cheap costs and automated procedures and hypothetically interrupt the budget, as we classify it, by dropping the addiction on central units.

The development of smart stuff has extremely preferred the impression of linking everyday stuff via the present systems. The melodramatic growth of smart devices headed towards the recognition of IoT [23]. The IoT is an uninterruptedly emergent network that unconventionally classifies and bonds data amongst exclusively addressable strategies [24].

Artificial Intelligence (AI) can help in numerous types of IoT data. AI helps in data preparation in which pools of data are defined and clean them which gives us the data lakes and dark data. Basically, dark data is the dusty data [25]. Gartner elaborates Information assets of an organization as a dark data. Organizations during their regular business activities gathered, process and manipulate data, but mainly, not beneficial for

other different purposes. Data lake is beneficial especially when you don't even know about the scope of data and uses because data lake is designed to retain all attributes. [26]. AI plays an important role in predictive and advanced analytics at this stage decision can be made based on data collection, data discovering and data analyzation [27].

We have to face many challenges while using AI with data. One of the biggest challenges is compatibility challenge because we have a collection of different parts of the system which is fundamentally different and complicated system which produces a very complex ecosystem. It generates many privacy, safety and security issues.

VII. IMPORTANCE OF CLOUD COMPUTING

Theory of grid computing has strong roots of cloud computing [28], It plays an important role to diminish computation cost and by gathering computing possessions update the elasticity and reliability. NIST elaborates that [29], A model for permitting appropriate, persistent on-demand access of a network to a combined lake of computer possessions that are configurable like networks, server, storage, application and services this can instantly provisioned and released with minimum service provider interaction and management effort."

Virtualization technology is used in cloud computing which takes a variety of altered categories of computing assets as distracted facilities to users. These cloud services are distributed into Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS).

Cloud computing has expectancy and connectivity issues because of the intrinsic distribution of resources which causes a problem for some specific users. In order to secure the access to resources without fatalities, recovery of the appliances and fault tolerance must be provided by the providers (service providers). Reduction of the fingerprint to associate big cloud data centers are one of the main challenges in the energy efficiency field. In detail, a number of companies encouraged their infrastructure where they managed it more definitely and economically. It also helps to appraise the organizations with optimize-power microchips that shrink power ingestion.

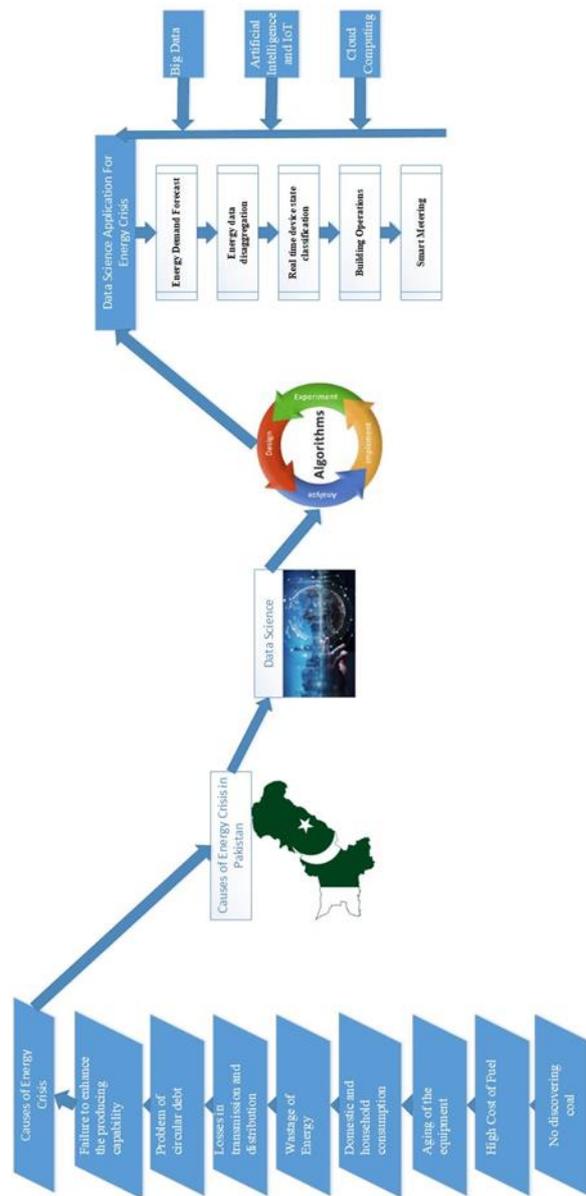


Figure 5 Flow chart of energy crisis solution in Pakistan

VIII. CONCLUSION

This study discovered and delivered comprehensive review affecting the existing position of the energy sector of Pakistan. it is a well-acknowledged circumstance that energy systems are becoming additional varied and unconventional with a diverse combination of consumption and generation technologies. The basic demand, for now, is that the administration, as well as the private establishments, essentially pay thoughtful consideration to this main problem. Struggles must be made to exploit the obtainable energy possessions, to discover fresh one maximum energy resources in order to overwhelm the hole between energy supply and demand. Numerous foreign establishments can also be accessed in this esteem. It is very important to manage and

construct the buildings in a sustainable way because building segment is a massive energy consumer.

On this review, the crucial purposes of this study were attained with analysis of Pakistan's energy sector with definite emphasis on Energy Sectors in Pakistan in the Section 2, comprehensive analysis of various causes of energy crisis in Pakistan in Section 3, review of effects of energy crisis in Pakistan in Section 4, comprehensive overview of Data Science and some frequently used algorithms in Section 5, provide important review of Data Science application for energy crisis in Section 6, provide brief description of Big Data in Section 7, Section 8 review the importance of Artificial Intelligence and IoT, Section 9 provides the review of cloud computing for energy management.

REFERENCES

- [1] Farhangi H. The path of the smart grid. *IEEE Power Energy Mag* 2010;8(1):18–28.
- [2] (<http://worldpopulationreview.com/countries/Pakistan-population/>) dated 28-02;2017.
- [3] Pakistan's Energy Sector. From crises to crises breaking the chain. Pakistan Institute of Development Economics Islamabad; 2012.
- [4] Hydrocarbon Development Institute of Pakistan. Pakistan energy yearbook 2001; January 2002.
- [5] Power Sector Institutions, Page 1; Policy for Development of Renewable Energy for Power Generation. Government of Pakistan; 2006.
- [6] World Times Publication "Causes of Energy Crises in Pakistan", 2016
- [7] ADI SHAMIR "Uncover teh Real Application of Data Science"
- [8] Sunil Ray "Types of Regression techniques"
- [9] Jain AK, Murty MN, Flynn PJ. Data clustering: a review. *ACM Comput Surv* 1999;31(3):264–323.
- [10] Miguel Molina-Solana, b, María Rosa, □, M. Dolores Ruiz, Juan Gómez-Romero, M.J. Martín-Bautista "Data science for building energy management: A review"
- [11] Yang G, Tumwesigye E, Cahill B, Menzel K. Using data mining in optimization of building energy consumption and thermal comfort management. In: Proceedings of the 2nd international conference on software engineering and data mining (SEDM), 2010, June 2010. pp. 434–9.
- [12] May-Ostendorp PT, Henze GP, Rajagopalan B, Corbin CD. Extraction of supervisory building control rules from model predictive control of windows in a mixed mode building. *J Build Perform Simul* 2013;6(3):199–219
- [13] Ahmed A, Ploennigs J, Gao Y, Menzel K. Analyse building performance data for energy-efficient building operation. In: Proceedings of the 26th international conference on managing IT in construction, Istanbul, Turkey, 2009.
- [14] Kusiak A, Li M, Tang F. Modeling and optimization of HVAC energy consumption. *Appl Energy* 2010;87(10):3092–102
- [15] Xiao F, Fan C. Data mining in building automation system for improving building operational performance. *Energy Build* 2014;75:109–18.
- [16] McDaniel P, McLaughlin S. Security and privacy challenges in the smart grid. *IEEE Secur Priv* 2009;7(3):75–7
- [17] Efthymiou C, Kalogridis G. Smart grid privacy via anonymization of smart metering data. In: 2010 Proceedings of the first IEEE international conference on smart grid communications (SmartGridComm). 2010. pp. 238–43.
- [18] European Commission Press Release. European commission and data industry launch 2.5 billion partnership to master big data. [Online; [accessed 15.12.14]
- [19] Schmarzo B. Big data: understanding how data powers big business. Wiley; 2013
- [20] Lesser A. How energy data will impact the smart grid, 2013. [Online; [accessed 13.04.15] (<http://research.gigaom.com/report/how-energy-data-will-impact-the-smart-grid/>)
- [21] IBM. Managing big data for smart grids and smart meters. White paper, IBM Software, 2012.
- [22] Rifkin J. The zero-marginal cost society: the internet of things, the collaborative commons, and the eclipse of capitalism. St. Martin's Press; 2014.
- [23] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari and M. Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 4, pp. 2347 - 2376, 2015
- [24] M. Razaque, M. Milojevic-Jevric, A. Palade and S. Clarke, "Middleware for Internet of Things: A Survey," *IEEE Internet of Things Journal*, vol. 3, no. 1, pp. 70 - 95, 2016.
- [25] Open Mind "Understanding Dark Data" 2015
- [26] Open Mind "Data Lake: An Opportunity or a Dream for Big Data?" 2015
- [27] Open Mind "Why IoT Needs AI" 2017
- [28] Buyya R, Broberg J, Goscinski A. editors. Cloud computing: principles and paradigms. Wiley, 2011.
- [29] National Institute of Standards and Technology (NIST). The NIST definition of Cloud Computing. (<http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>). [accessed 07.12.16].
- [30] Baloch MH, Kaloi GS, Memon ZA. Current scenario of the wind energy in Pakistan challenges and future perspectives: a case study. *Energy Rep* 2016; 2:201–10.
- [31] Siddique S, Wazir R. A review of the wind power developments in Pakistan. *Renew Sustain Energy Rev* 2017; 1–61
- [32] Shami SH, Ahmad J, Zafar R, Haris M, Bashir S. Evaluating wind energy potential in Pakistan's three provinces, with proposal for integration into national power grid. *Renew Sustain Energy Rev* 2016; 53:408–21