TEST Engineering & Management

January – February 2020
ISSN: 0193-4120 Page No. 15205 - 15212

Virtual Assistant for Farmers using Machine Learning

R. Subhashini[1],TadikondaPrudhvi Siva Sai Krishna[2], TatiSudheer[2], C.Kavitha[3], E.Brumancia[4], Y. Bevish Jinila[5]

[1]Professor, School of Computing,Sathyabama Institute of Science and Technology, Chennai, India
[2]Student, School of Computing,Sathyabama Institute of Science and Technology, Chennai, India
[3]Asst. Professor, Dept., of Mathematics, Sathyabama Institute of Science and Technology, Chennai, India
[4]Asst. Professor, School of Computing,Sathyabama Institute of Science and Technology, Chennai, India
[5]Assoc. Professor, School of Computing,Sathyabama Institute of Science and Technology, Chennai, India

Article Info
Volume 82
Page Number: 15205 - 15212
Publication Issue:
January-February 2020

Article History
Article Received*: 18 May 2019*
Revised*: 14 July 2019*
Accepted*: 22 December 2019*
Publication*: 28 February 2020*

Abstract:
Every one works hard just to earn some food and survive but these days it has been very hard to produce sufficient food for the mammoth population across the world. This was due to lack of knowledge about modern practices and technologies in agriculture which will boost the fertility of land and food production. The aim of the proposed method is to provide a virtual Assistant that enables the user to know about modern practices,efficient techniques and appropriate amount of fertilizers and water for a particular crop which functions either by text input or voice commands. The text or voice commands are converted to machine understandable language by semantic parsing technique later on the mechanism forwards to similarity function which helps in finding out keywords and probability of them by Naive Bayes algorithm and gives the output in voice format. The proposed virtual assistant operates in regional language which stands apart from the existing models.

Keywords : Virtual Assistant, Machine Learning, Farmers, Naïve Bayes algorithm, Chatbot

1. Introduction

A menial helper is a product operator that can perform errands or administrations for a person. In some cases the expression "chatbot" is utilized to allude to remote helpers for the most part or explicitly those gotten to by online visit (or now and again online talk programs that are for excitement and not valuable purposes).Apple and Google have huge introduced bases of clients on cell phones. Microsoft has a huge introduced base of Windows-based PCs, cell phones and shrewd speakers. Alexa has a vast introduce base for savvy speakers. Data and correspondence innovation in agribusiness otherwise called e-horticulture is creating and applying inventive approaches to utilize Information and Communication Technology in the rustic space, with the essential spotlight on farming. Data innovation offers a wide scope of answers for some horticultural difficulties. We as a whole know soil and water is the key component of horticulture, yet constant trimming changes soil physiochemical parameters, proteins and microorganism which causes replant issue. So ranchers are driven leave the cultivable land into non-cultivable land. The dirt test is done in such land and the data is put away in the database, last investigation give data about the best harvest to be planted dependent on the ebb and flow showcase request which return more benefit to the farmers .The proposed framework helps the ranchers by giving important solutions to their questions about horticulture.

For example, manure the board,sort of soil essential for compelling plant development ,climatic conditions for the estate,water level required for plant to develop with proper nutritive esteem .The whole procedure should be possible in any language over the globe dependent on their comfort. The task of menial helper in provincial language is a vital component.In all the previously mentioned and

January – February 2020
ISSN: 0193-4120 Page No. 15205 - 15212

existing frameworks. Their task is done just in English language yet the proposed has the ability to work in every one of the dialects over the globe which incorporates English additionally.Moreover, there are no remote helpers allotted or created to work for a specific field however this venture has a specific field required to give benefits in horticulture. There are numerous applications in agribusiness yet they work just in English, which the vast majority of the ranchers didn't know off particularly in Asian nations. Thus, this proposed framework would be an incredible help for ranchers absent much trouble just as effectively justifiable.The most concerning issue for ranchers is absence of training in Asian nations and a few people don't have a clue how to peruse too .This issue is dealt with by this proposed framework which has voice directions for contribution just as yield,which is a champion element of this venture.As this task expects to help the ranchers by giving better learning about inventive and compelling practices and present day mechanical strategies and frameworks which will be valuable for high creation rate.

2. Related Work

As headways have occurred, a huge amount of studies and research have happened on remote helpers.The accompanying papers were contemplated and profoundly examined so as to build up our task. The papers which are referenced beneath are distributed as a piece of prestigious diaries and were genuinely useful in understanding the possibility of Virtual Assistants.

One of the objectives of Artificial knowledge (AI) is the acknowledgment of normal exchange among people and machines [5]. As of late, the exchange frameworks, otherwise called intelligent conversational frameworks are the quickest developing territory in AI. Numerous organizations have utilized the discourse frameworks innovation to build up different sorts of Virtual Personal Assistants(VPAs) in light of their applications and territories, for example, Microsoft's Cortana,Google Assistant, Apple's Siri, Amazon Alexa and Facebook's. Be that as it may, in this proposition, we have utilized the multi-modular discourse frameworks which process at least two consolidated client input modes, for example, discourse, picture, video, contact, manual signals, look, and head and body development so as to plan the Next Generation of VPAs display. The new model of VPAs will be utilized to expand the association among people and the machines by utilizing diverse innovations, for example, signal acknowledgment, picture/video acknowledgment, discourse acknowledgment, the immense exchange and conversational learning base, and the general information base. In addition, the new VPAs framework can be utilized in other diverse zones of utilizations, including instruction help, therapeutic help, mechanical autonomy and vehicles, incapacities frameworks, home robotization, and security get to control.

On-going advancements in brilliant aides and keen home mechanization are of late pulling in the intrigue and interest of purchasers and scientists [6]. Discourse empowered remote helpers (frequently named shrewd speakers) offer a wide assortment of system situated administrations and, at times, can associate with keen conditions, along these lines improving them with new and compelling UIs. Notwithstanding, such gadgets likewise uncover new needs and a few shortcomings. Specifically, they speak to anonymous and daze associates, unfit to demonstrate a face, and accordingly a feeling, and unfit to 'see' the client. As an outcome, the connection is impeded and, sometimes, inadequate. Additionally, the vast majority of those gadgets intensely depend on cloud-based administrations, along these lines transmitting conceivably touchy information to remote servers. To defeat such issues, in this paper we consolidate the absolute most exceptional procedures in PC vision, profound learning, discourse age and acknowledgment, and man-made consciousness, into a remote helper engineering for savvy home robotization frameworks. The proposed collaborator is compelling and asset productive, intelligent and adjustable, and the acknowledged model keeps running on an ease, small sized, Raspberry PI 3 gadget. For testing purposes, the framework was integrated with an open source home automation condition and kept running for a few days, while individuals are urged to associate with it, and ended up being precise, solid and engaging.

J. Jiang et al., [7] proposed a new video likeness compute demonstrate utilizing video time thickness work (VTDF) and dynamic programming. Initially, we utilize VTDF to portray the thickness of video

January – February 2020
ISSN: 0193-4120 Page No. 15205 - 15212

exercises in time space by computing the between edge shared data. Second, a fleeting allotment arrangement is connected to partition every video grouping into equi-sized worldly sections. Third, another VTDFbased similitude measure utilizing relationship is determined to gauge the closeness between two worldly portions. Fourth, unique writing computer programs are then created to locate the ideal non-direct mapping between two video groupings. Another standardized similitude measure work brushing both visual attributes and transient data together is to assess the semantic likeness of two video successions. Test results demonstrate that the proposed estimation show is viable to investigate the semantic comparability of video arrangements.

Remote Sensor Networkis new innovation to world and nation similar to India where it can utilized in Agriculture Sector in India for expanding capitulate by giving early forecast of plant infections and pest [8].This can be occur by taking crude information from field where WSN organize is introduce and with fitting proper AI display for this information to get anticipated output .[9] gives a plan to how to convey WSN on field and how Machine learning model is fitted for expectation of vermin/sicknesses utilizing Naive Bayes Kernel Algorithm. Clever checking framework is required in the field which should comprises of cutting edge sensor innovation with extra data handling innovation to get ongoing information from the field in precise and quantitative way to take care of issue of plant illnesses and nuisance related issues. [10]With the assistance of Wireless sensor arrange framework choice can be set aside a few minutes/spot and contribution at small scale climatic dimensions. Harvest misfortunes because of irritation and ailments are progressively impressive specifically in semi-dry conditions where they cause direct harm to harvest and yield where Temperature, stickiness, precipitation, leaf wetness, windspeed and soil supplements are fundamental parameters affecting the bug/sicknesses frequency. There is developing measure of Data accessible from numerous assets that can be utilized for picking up information in agribusiness part .In request to anticipate helpful data just as to comprehend the patterns, the information need to broke down and utilized for taking choice where calculation method should be intended to perform separate work and to substitute for rancher requirements .Such information can be utilized for beneficial basic leadership where AI system are connected. A proficient model is required for motivation behind estimating and identifying the bug/ailments precisely. The proficient model strategy and sensor information have been utilized to comprehend and evaluate concealed relationship bother/sicknesses with on-going sensor information. For that we accompany Naive Bayes Kernel show where we are understanding connection design between on-going information and existing dataset II. Writing SURVEY Many methodologies are agreed for anticipating the agribusiness crop results which is given in this area Pratheepa.et al[1] has created Classification display for admonishing about the cotton crop bug and furthermore recognizing the components affecting the irritation populace thickness which will additionally assist rancher with applying vermin control procedures on time to decrease crop misfortune.

3. Proposed system

The importance of Machine Learning is improved over contribution Analysis. In this research train machines to scenarios, recognize and tag events, or forecast a value in the present or future, data science is of the quintessence. It is vital to study the underlying data and model it by choosingasuitable algorithm to approach any such use case. The various control parameters of the algorithm need to be pinched to fit the data set. As learn with time how to define a result, the developed application progresses and becomes more efficient in solving the problem. Here we are going to give an input for queries in voice input format .Then we here are using various south Indian languages for voice input .This south Indian languages adding in our project by using Google API that it will convert the regional language format to English format .This is will be converted to string , by using similarity function we are going to get the keywords .Then parsing the keywords to get their base meaning .After that by using naive bayes algorithm [11]finding the probability for the each question.Question with highest probability will return the appropriate answer for the question in the database. If the solution is not found (means less than threshold value) then searching the query in the Google and display the answer in their given input regional language.

January – February 2020
ISSN: 0193-4120 Page No. 15205 - 15212

3.1 Architecture

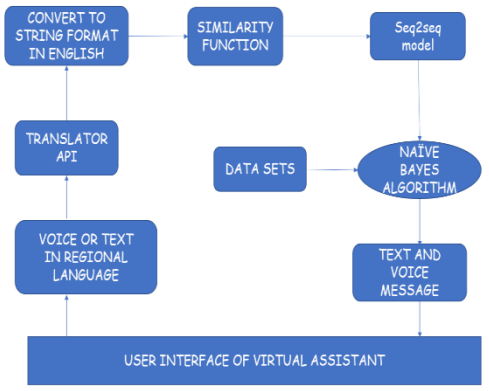


Fig.1. Architecture Diagram

The architecture of virtual assistant for farmers starts with the input given by the user. If the input given by the user is in English, then the similarity function takes place. By Similarity function the keywords in the input are matched with the keywords present in the data sets by sequence to sequence model. Naive bayes algorithm helps in finding out the probability of keywords when compared with the keywords in data sets. The maximum probability is displayed as the answer if the maximum probability value is greater than the threshold value. If it is lower than the threshold value the user will be directed to Google to search for results. The output is displayed in the form of text or voice commands .If the input of the user in regional language other than English the Google translator API helps to translate the voice commands in to English. Later, then the similarity function takes place. By Similarity function the keywords in the input are matched with the keywords present in the data sets by sequence to sequence model. Naive Bayes algorithm helps in finding out the probability of keywords when compared with the keywords in data sets. The maximum probability is displayed as the answer if the maximum probability value is greater than the threshold value. If it is lower than the threshold value the user will be directed to Google to search for results. The output is displayed in the form of text or voice commands .The text to speech mechanism is used to prompt the answers to the query of user.

3.2 Algorithm Used

Naive Bayes classifiers are extremely adaptable, requiring a variety of parameters directly in the number of factors (highlights/indicators) in a learning concern. Most extreme probability preparing should be feasible by assessing a shut structure articulation, which takes straight time, as a substitute of by costly iterative estimation as exploited for some different sorts of classifiers. In the insights and software engineering writing, credulous Bayes models are known under a variety of names, together with basic Bayes and autonomy Bayes.All the names reference the utilization of Bayes' hypothesis in the classifier's choice principle except innocent Bayes isn't (really) a Bayesian technique. Innocent Bayes is an elementary strategy for constructingclassifiers models that dole out class marks to concern occurrences, spoke to as vectors of highlight esteems, everywhere the class names are drawn from a few limited set. There is definitely not a solitary calculation for preparing such classifiers, yet a group of calculations dependent on a typical guideline: all Naive Bayes classifiers expect that the estimation of a specific element is autonomous of the estimation of some other component, given the class variable. For instance, a natural product might be viewed as an apple in the event that it is red, round, and around 10 cm in measurement.

$$p(C_k \mid \mathbf{x}) = \frac{p(C_k)\, p(\mathbf{x} \mid C_k)}{p(\mathbf{x})}$$

4 Methodology

i. Google voice input:

Android telephones have the astonishing capacity to decipher your articulations as content. It works nearly just as PC correspondence in sci-fi films; however it won't discover the order to crush Alderaan.

Semantic similitude is a measurement characterized over a lot of archives or terms, where separate between them depends on the resemblance of their significance or semantic substance rather than closeness which can be assessed in regards to their linguistic portrayal (for example their string group). These are scientific devices used to gauge the quality of the semantic connection between units of language, ideas or occasions, through a numerical depiction got by the examination of data supporting their importance or portraying their temperament. The term semantic comparability is frequently mistaken for semantic relatedness. Semantic relatedness incorporates any connection between two terms, while semantic comparability just incorporates "is a" relations. For instance, "vehicle" is like "transport", but on the other hand is identified with "street" and "driving". Computationally, semantic closeness can be assessed by characterizing a topological similitude, by utilizing ontologies to characterize the separation between terms/ideas. For instance, a guileless measurement for the correlation of ideas requested in an in part requested set and

January – February 2020
ISSN: 0193-4120 Page No. 15205 - 15212

spoke to as hubs of a coordinated non-cyclic chart (e.g., a scientific categorization), would be the most limited way connecting the two idea hubs. In light of content investigations, semantic relatedness between units of language (e.g., words, sentences) can likewise be evaluated utilizing factual methods, for example, a vector space model to associate words and printed settings from an appropriate content corpus. The idea of semantic similitude is more explicit than semantic relatedness, as the last incorporates ideas as antonym and metonymy, while comparability does not.

- Activating Voice Input
- Dictating text

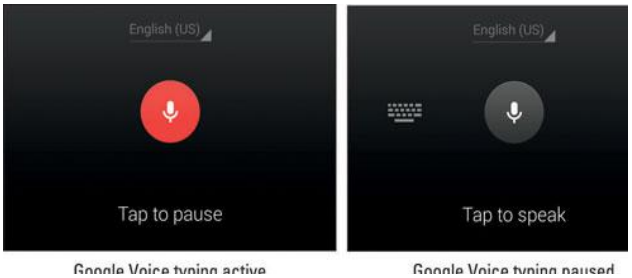


Fig.2. Google voice input

ii. Google Translator API :

The Google Translation API can progressively decipher message between a huge numbers of language sets. The Cloud Translation API gives sites and projects a chance to incorporate with the interpretation administration automatically. The Cloud Translation API is a piece of the bigger Cloud Machine Learning API family. Google Translate is a multilingual machine interpretation administration created by Google, to translate content. It offers a site interface, versatile applications for Android and IOS, and an API that assists designers to manufacture program expansions and programming applications.

iii. Similarity Function :

Semantic similitude is a measurement characterized over a lot of archives or terms, where separate between them depends on the resemblance of their significance or semantic substance rather than closeness which can be assessed in regards to their linguistic portrayal (for example their string group). These are scientific devices used to gauge the quality of the semantic connection between units of language, ideas or occasions, through a numerical depiction got by the examination of data supporting their importance or portraying their temperament. The term semantic comparability is frequently mistaken for semantic relatedness. Semantic relatedness incorporates any connection between two terms, while semantic comparability just incorporates "is a" relations. For instance, "vehicle" is like "transport", but on the other hand is identified with "street" and "driving". Computationally, semantic closeness can be assessed by characterizing a topological similitude, by utilizing ontologies to characterize the separation between terms/ideas. For instance, a guileless measurement for the correlation of ideas requested in an in part requested set and spoke to as hubs of a coordinated non-cyclic chart (e.g., a scientific categorization), would be the most limited way connecting the two idea hubs. In light of content investigations, semantic relatedness between units of language (e.g., words, sentences) can likewise be evaluated utilizing factual methods, for example, a vector space model to associate words and printed settings from an appropriate content corpus. The idea of semantic similitude is more explicit than semantic relatedness, as the last incorporates ideas as antonym and metonymy, while comparability does not.

iv. Naive Bayes Algorithm :

Naive Bayes classifiers are extremely adaptable, requiring various parameters straight in the quantity of factors (highlights/indicators) in a learning issue. Most extreme probability preparing should be possible by assessing a shut structure articulation, which takes straight time, instead of by costly iterative estimation as utilized for some different sorts of classifiers. In the insights and software engineering writing, credulous Bayes models are known under an assortment of names, including basic Bayes and autonomy Bayes.All these names reference the utilization of Bayes' hypothesis in the classifier's choice principle, however innocent Bayes isn't (really) a Bayesian technique. Innocent Bayes is a basic strategy for building classifiers: models that dole out class marks to issue occurrences, spoke to as vectors of highlight esteems, where the class names are drawn from some limited set. There is definitely not a solitary calculation for preparing such classifiers, yet a group of calculations dependent on a typical guideline: all Navie Bayes classifiers expect that the estimation of a specific element is autonomous of the estimation of some other component, given the class variable. For instance, a natural product might be viewed as an apple in the event that it is red, round, and around 10 cm in measurement.

$$p(C_k \mid \mathbf{x}) = \frac{p(C_k)\, p(\mathbf{x} \mid C_k)}{p(\mathbf{x})}$$

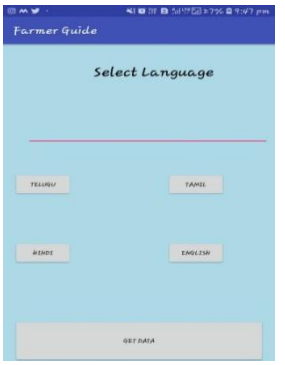
January – February 2020
ISSN: 0193-4120 Page No. 15205 - 15212

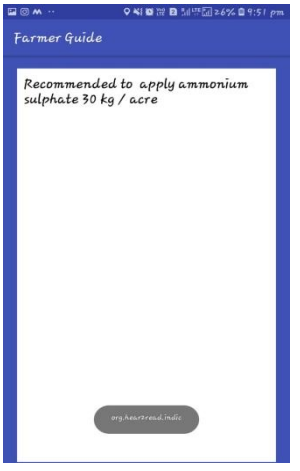
v. Google text to speech output:

Cloud Text-to-Speech API enables engineers to make common sounding, manufactured human discourse as playable sound. You can utilize the sound information documents you make utilizing Cloud Text-to-Speech API to control your applications or expand media like recordings or sound chronicles (in consistence with the Google Cloud Platform Terms of Service incorporating consistence with all material law). The Text-to-Speech API changes over content or Speech Synthesis Mark up Language (SSML) contribution to sound information like MP3 or LINEAR16 (the encoding utilized in WAV documents). This archive is a manual for the basic ideas of utilizing the Cloud Text-to-Speech API. Prior to jumping into the API itself, audit the speedy begin. The Text-to-Speech API is perfect for any application that plays sound of human discourse to clients. It enables you to change over subjective strings, words, and sentences into the sound of individual talking similar things. Envision that you have a voice associate application that gives characteristic language input to your clients as playable sound records. Your application may make a move and after that give human discourse as criticism to the client.

5. RESULTS AND DISCUSSION

The proposed system helps the farmers to find the current demand of crop cultivated in their own land, which result in awareness about the crop rotation and farmer's growth. There is anessential of identifying optimum matching crops in farming. The conventional ways of farming cannot sustain farmer to attain the necessary yield. Therefore a digitized method or system should be designed to help farmer to augment the yield in their farms. Since India is an Agrarian nation, the system will provide a generalized solution for most of the problems in farming related to cultivation of crops. The main reason of developing such system is to assist farmers to augment the productivity of their fields and to enhance the gross domestic product and diminishthe poverty in India.





TEST Engineering & Management

January – February 2020
ISSN: 0193-4120 Page No. 15205 - 15212

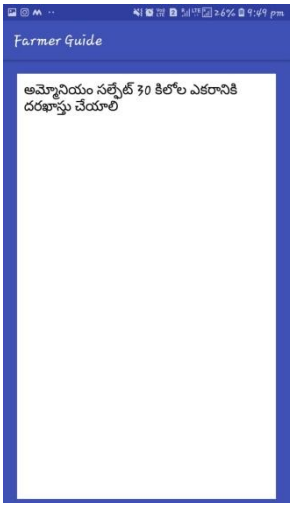


Fig.3. Output

6. Conclusion and Future Work

The proposed system helps farmers to query about the agriculture, get the response in text as well as speech. It receives and also gives the responses in their regional language.The proposed system provides better user experience by natural language processing. The future enhancement can be done by in predicting the forthcoming data of price, so that they can strategy their activities. Thus, there are a lot of things that can be upgraded for better user experience.

Acknowledgement

We wish to acknowledge the Department of Science and Technology, India and School of Computing, Sathyabama Institute of Science and Technology, Chennai for providing the facilities to do the research under the DST-FIST Grant Project No.SR/FST/ETI-364/2014.

References

1. V. Këpuska and G. Bohouta, "Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home)," *2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC)*, Las Vegas, NV, 2018.
2. G. Iannizzotto, L. L. Bello, A. Nucita and G. M. Grasso, "A Vision and Speech Enabled, Customizable, Virtual Assistant for Smart Environments," *2018 11th International Conference on Human System Interaction (HSI)*, Gdansk, 2018.
3. J. Jiang, X. Zhang and A. C. Loui, "A new video similarity measure model based on video time density function and dynamic programming," 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, 2011.
4. H. Wani and N. Ashtankar, "An appropriate model predicting pest/diseases of crops using machine learning algorithms," 2017 4th International Conference on Advanced computing and Communication Systems(ICACCS),Coimbatore,2017
5. V. Këpuska and G. Bohouta, "Next-generation of virtual personal assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home)," 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC), Las

January – February 2020
ISSN: 0193-4120 Page No. 15205 - 15212

Vegas, NV, 2018, pp. 99-103. doi: 10.1109/CCWC.2018.8301638

6. G. Iannizzotto, L. L. Bello, A. Nucita and G. M. Grasso, "A Vision and Speech Enabled, Customizable, Virtual Assistant for Smart Environments," 2018 11th International Conference on Human System Interaction (HSI), Gdansk, 2018, pp. 50-56. doi: 10.1109/HSI.2018.8431232
7. J. Jiang, X. Zhang and A. C. Loui, "A new video similarity measure model based on video time density function and dynamic programming," 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, 2011, pp. 1201-1204. doi: 10.1109/ICASSP.2011.5946625
8. H. Wani and N. Ashtankar, "An appropriate model predicting pest/diseases of crops using machine learning algorithms," 2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, 2017, pp. 1-4.
9. 9.V.VijeyaKaveri, V.Maheswari , “Mining Social Data to Identifying User Behavior in Medhelp Forum on Health-Related Topics “, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-2S3, July 2019, PP.No.1306-1310 .
10. 10.P.Ajitha, A. Sivasangari, K.Indira,"Predictive Inter and Intra Parking System",International Journal of Engineering and Advanced Technology(IJEAT)ISSN: 2249 – 8958, Volume-8, Issue-2S, December 2018.
11. 11.Saranya, E., B. Baron Sam, and R. Sethuraman. "Speech to text user assistive agent system for impaired person." 2017 IEEE International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM).IEEE, 2017.