Cloud based Cognitive and Predictive Real-time Personal Advisor for an Agriculturist

Shubhradeep Nandi

Data Scientist and Research Scholar IBM
E-mail: shunandi@gmail.com

Abstract—*Today an agriculturist is at the mercy of nature, unskilled local help, and not so localized government guidance.*

The real-time knowledge that an agriculturist needs in-order to fight against a possible pest infestation, or, to find alternatives to overcome an approaching drought, or, to identify the right selling price for his/her crop is not available anywhere.

Hence, all these uncertainties finally leads an agriculturist to
a)   Either loose his produce or yield low quality crops,
b)   Not get the right price for his invested money and hard work,
c)   Force him/her for a distress sale at a very low price.
This thus leads to piling up of financial burden on him. This financial burden becomes larger and larger year on year leading him/her to change his/her profession or worse commit suicide.

According to an official statistics each day over 2000 farmer are leaving agriculture in India and between the years 2010 – 2014 more than sixty seven thousand farmers have committed suicide [1]. This research paper is a humble effort to help "those hands that feed us" to live a better life using advances in data science and allied machine learning technologies.

1. INTRODUCTION

In India agriculture and allied work composes of about 50% workforce but accounts to only 12.53% of the GDP (Gross Domestic Product) making it one of the lowest valued job.

According to a study done by CSDS (Centre for studying of developing societies) 76% of the farmers in India want to give up farming. The few primary reasons behind this decision are poor income, bleak future and piling up of financial burden. Today there is none real-time, localized, and personalized farming guideline available that can help an agriculturist to deal with various farming related challenges.

2. PERSONAL ADVISOR SYSTEM

A personal advisor system is one that can exhibit deeper knowledge in a specific area. The role of this system is to mentor an individual with recommendations based on real-time data.

The role of a personal advisor system is to get into a conversation with the system's user, understand the problem scenario at hand based on the spoken, written or visual information shared by the user. Once the problem scenario is completely understood it does its due diligence and comes back with the most probable solution. A personal advisor has access to all relevant systems from which it can gather information and sufficiently advise the user to overcome a problem along with the logical solution steps to be followed.

3. COGNITIVE COMPUTING

A Cognitive computing system is capable to put content in context of conversation, provide probabilistic responses, and then support them with document based evidence. They can also do a rapid exploration like discovering a needle in bush. They are also capable of identifying new patterns and deliver insights from them.

Decades of research have made cognitive system capable of simulating a human brain. They now can help solve the most complex and pressing problems by diving into the complexity of big data. They can also utilize the strong capabilities of natural language processing and machine learning.

We have explored three key capabilities of cognitive computing systems that is Engagement, Decision and Discovery as shown in Figure 1. We have focused on how existing cognitive computing services can help us device a system that is easy to use, can communicate in the same language and linguistics of an agriculturist and provide real-time insights, suggestions and recommendations based on various knowledge sources [2].

**Figure 1. Capabilities of Cognitive System**
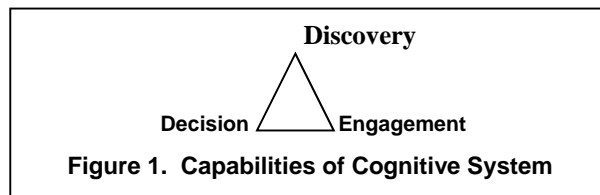
4. CLOUD COMPUTING

Cloud computing is Internet-based way of computing which provides a way of shared computer processing and data on demand as shown in Figure 2. It is a way for allowing ubiquitous, on-need access to a shared pool of configurable

resources. It can be rapidly provisioned and released with minimal management effort. Cloud computing and storage solutions provide users with various capabilities to store and process their data in centers[3] that may be located far from the user. Cloud computing relies on sharing of resources to achieve coherence and economy of scale.

With multiple methods and models of security, visibility and access to external systems and services across the globe, elastic scalability in terms of computation, high availability without failure, and complete commitment without owning any infrastructure has made cloud computing the most preferable platform for large volume computation. One of the progressive system that can be used to attain this is IBM Bluemix [3]

Figure 2. Cloud computing System
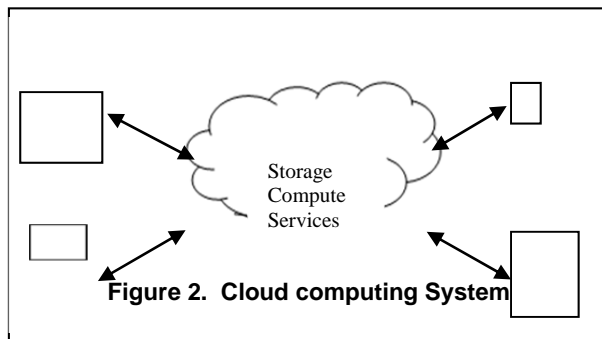
5.   COGNITIVE   AND   PREDICTIVE   PERSONAL ADVISOR ON CLOUD

Cognitive and predictive personal advisor system is one that can enter into a conversation with the agriculturist and provide him/her with relevant solutions.

The complete process involves the following steps:-

1.   **Understand the query of the agriculturist in his/her native language.**
This is a cognitive capability and is currently supported by multiple natural language processing systems as a service. One such system is IBM Watson which has a robust

Speech to Text technology. From Indian context Sampark [4] and CDAC [5] has developed a speech recognition module that supports multiple Indian languages using complex data processing techniques as shown in Figure 3.

Audio → Speech Recognition (acoustic + language + domain model) → text alignment (transcription) → Aligned text
Figure 3. Process involved in a speech to text NLP system.

2.   **Understand the intention and underlying entities of the discussion.**
This is a natural language processing characteristics in which the spoken query that is converted into text is classified to map the right answer based on the intention and underlying entities.

Once such API is conversation which is part of IBM Watson that can in its most simple form help classify variety of random queries into appropriate questions.

3.   **Adapts to the context of the discussion.**
This is a natural language processing characteristic in which the context of the conversation held for each user and propagated back and forth in order to give a seamless interaction. IBM Watson Conversation API is capable of maintaining this interaction context and with unique context identifier can maintain the context.

4.   **Cognitively decides from where to fetch the data, does real-time query, and fetches it.**
This is a cognitive characteristic in which the system is capable of identifying the appropriate data source and can decisively pull out the right data as shown in Figure 4. It can also take help of external application programming interfaces to get the data. IBM Watson has the capability to understand and harness data both from structured and unstructured dark data sources. The data in dark sources can be in the format of image, video, pdfs, documents and/or html.

Figure 4. Cognitive data source resolution
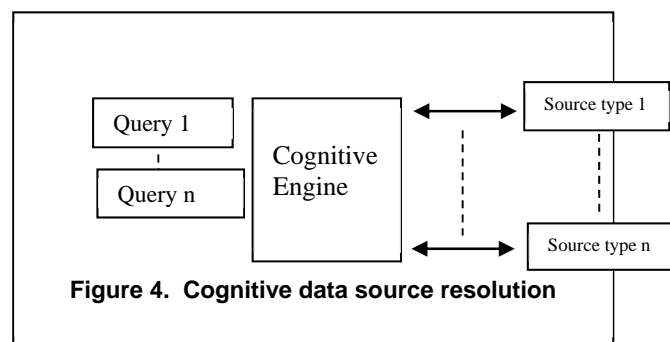
5.   **Predicts what to recommend based on various decision making parameters and data points using machine learning.**
This is a machine learning problem where based on the identified query it can first be classified as a supervised regression or classification problem or an unsupervised clustering problem. Once the algorithm decision is made, the data is retrieved then the best fit model is identified and the prediction is done to identify the right resolution set.

6.   **Does a tradeoff based on the possibilities.**
In open end queries during a conversation the resolution set can be large with multiple possibilities. But based on the user's profile, prioritized needs, and localized attributes a tradeoff is done in order to identify the appropriate solution. IBM Watson has got a dedicated tradeoff service to help with the tradeoff decision management.

7.   **Talks back to the agriculturist with the resolution in his/her local language.**
This is a cognitive capability and is currently supported by multiple natural language processing systems as a service that used complex data processing techniques as shown in Figure 5. One such system is IBM Watson which has a robust Text to

Speech technology with expressive SSML and voice transformation. From Indian language context CDAC [5] has also developed a text to speech module that supports multiple Indian languages.

Raw Text → Sentence Segmentation → tokenization → Wave concatenation and Smoothening →Normalization →Suprasegmental Processing →Synthesized Speech

Figure 5. Process involved in a text to speech

---

6. CONSUMABLE BENEFITS FOR AN AGRICULTURIST

An agriculturist now just needs a phone or any transmission device via which he/she can either call, message, chat, text or send multimedia data to this personal advisor system. In response he/she will receive answers to all the agriculture related queries. In the below section a few queries and how they will be addressed is mentioned.

6.1 The monsoon schedule

An agriculturist can ask when will be the monsoon this year. To answer such a query the basis would not only be the monsoon guidance given by meteorological department, but will also include multiple other parameters like prediction based on historical data, global warming index, pollution index, underground water availability and consumption patterns etc. A comprehensive recommendation will be provided on the monsoon schedule.

6.2 The right crop to sow

An agriculturist can ask which would be the right crop to sow this season. To answer such a query the basis would include soil, climate, and water availability data of the locality, historical data of the crops that can be grown in that area, demand and price forecasting for the crop based on retail need[6], normalized personal financial risk that the agriculturist is capable of taking and available loan sources, the manpower cost and availability data, the nearest market that has demand for this crop, storage facilities in the region etc. based on all these parameters a comprehensive recommendation will be made.

6.3 The disease in crop

An agriculturist can ask what disease his crop is having by sending a set of images or by describing the symptoms. Once sufficient data is available from conversation with the agriculturist. The data will be used to cross validate across real-time sources and dark data sources to find out what disease is the crop having. Machine learning algorithms and cognitive visual recognition services from IBM Watson can help classify the crop/plant disease.

Occasionally a manual intervention will also be required with the help of a remote specialist/scientist who can take into consideration the data shared by agriculturist and classify the disease. This manual intervention is further added to the cognitive system as a measure of learning.

6.4 The selling price

An agriculturist can ask at what price he should sell his produce. The cognitive and machine learning system in-order to answer such questions would not only consider the market price but also consider the cost an individuals has incurred, the loan repayment if taken, transportation cost, personal expenses, investments and savings for the unforeseen and other allied expenses until next harvest.

This price decision based on multiple such parameter in a cognitive and regressive manner can help a farmer stabilize his financial background and sell his produce at the right price.

One major benefit in this approach is that the agriculturist is assured that he is not being fooled by a middle men who generally makes most of the profit by creating a wrong price perception and forces the farmer for a distress sale.

6.5 The pesticide and fertilizer to be used

An agriculturist can occasionally ask what pesticide and fertilizer he/she should use. The cognitive and machine learning algorithm behind the personal advisor will not only considers what is cheap and easily available but would also consider its effect on nature and human health. Lot of research and advances have happened in the field of organic fertilizer, bio manure, and nature friendly pesticides which are green alternatives available at low cost.

Taking into consideration the price, personal financial capability of an agriculturist, effects on consumer, cascaded effect on nature and allied other possibilities the cognitive machine learning algorithm does a tradeoff to identify the best possibility.

Once the right pesticide and fertilizer is identified, the personal advisor system finds the nearest location of the user where it is available and based on the users consent can help him/her procure and also get it delivered to his farm.

7. CONCLUSION

Agriculture sector is the largest employment generator yet the lowest GDP contributor. Apart from that day in and day out we can hear of farmer suicide incidents where the farmer killed himself and/or also his entire family because of piling up of loans and huge financial burden, low yield, not getting the right price or a bad crop.

Analyzing the historical data we have found that most of the agriculturist still use older methodologies of farming and depend only on historical data or generalized predictions. We have also found that agriculture is still not planned on real-

time data and no retrospective efforts are either taken up by the agriculturist, or village panchayat or any supportive government body to make things better. The apathy towards agriculture is not limited here, the people who are appointed by government to help and advice the farmers are still largely unskilled or not up-to-date with latest advances in agricultural sciences and farming techniques.

Technology like cloud, cognitive and predictive computing can help collate the right data in an unbiased manner, understand the localized needs, instantly connect the agriculturist to the right expert available at far off location, give an agriculturist real-time advice and help the him/her to both plan on short-term needs and long-term goals.

Such systems can also take care of urgent needs in time of distress by reaching out for right help, and guide the agriculturist towards right direction. It can also act as one stop mentor to identify the right place to buy agricultural needs, right place to sell the produce, right price to pitch while selling the produce.

Thus, the cloud based cognitive and predictive personal advisor can become the most trustable friend and mentor of an agriculturist. It ideally can guide and make agriculture profitable again for not only the middle man or the retailer but also for the hands that produce them.

This is my humble effort as a Data scientist to identify the right approach to help agriculturists by stopping them from either committing suicides or leaving the profession because we still need food to survive and cannot eat money.

Lastly, it is my sincere appeal that lets see that this research does not end up as another pile of paper. Please join me to make it a reality so that we can empower our agriculturist brothers and sisters to make agriculture profitable again.

8.   ACKNOWLEDGEMENTS

REFERENCES

[1] https://en.wikipedia.org/wiki/Farmers%27_suicides_in_India

[2] A. Lally, J. M. Prager, M. C. McCord, B. K. Boguraev, S. Patwardhan, J. Fan, P. Fodor, J. Chu-Carroll: Question analysis: How Watson reads a clue. IBM Journal of Research and Development 56(3.4), 2--1, IBM, (2012)

[3] Shubhradeep Nandi: Methodical Evaluation and Exploration of Bluemix - A Versatile Platform as a Service (Xpaas) that is built to Disrupt. , Indian Technology Congress, (2015).

[4] http://ilmt.tdil-dc.gov.in/sampark/web/index.php/content.

[5] http://cdac.in/index.aspx?id=mc_st_speech_technology.

[6] Shubhradeep Nandi: Analytical Prescription for Unorganized Retail in India. , Advances in Economics and Business Management (AEBM), Online ISSN: 2394-1553; Volume 2, Number 3; January-March 2015 pp. 241-245, Print ISSN: 2394-1545 9th International Conference, MBFEMT-2015, Jawaharlal Nehru University (2015).