KrishiEkta: Integrated Knowledge and Information Distribution

System for Indian Agriculture

Rakshit Agrawal  
CSC India Pvt. Ltd.  
Noida  
rakshit_agrawal@outlook.com

Mridu Atray  
Indian Institute of Technology  
Delhi  
mriduatray@outlook.com

S. Krishna Sundari  
Jaypee Institute of Information Technology, Noida  
krishna.sundari@jiit.ac.in

Abstract
KrishiEkta is an integrated knowledge distribution system for Indian Agriculture. KrishiEkta addresses the inadequate access of important information to various stakeholders in Indian agriculture. Access of rich internet platforms is not available to various farmers in small rural parts of the country. KrishiEkta system collects information from multiple sources through the internet on different categories falling under the theme of Agriculture. It then organizes this data into categorical models. This information with categorical information is the prepared for distribution in four different formats: SMS, tweet, Facebook post and digests e-mail. SMS are the preferred medium of information distribution among small farmers. Whereas many educated stakeholders use e-mails and social networking services. Summarized information with relevant links and identifiers per format are finally distributed to the subscribed users through the Distribution Engine.

Parsing Engine
- Fetch new news items from known sources
- Crawl through source web pages for article collection
- Parse through web pages for article data collection
- Push all article data and metadata into the database

Categorization Engine
- Assign categories to articles based on relevance
- Use article vocabulary to identify categories
- Use additional category bias based on article source
- Use associated tags/keywords as additional parameter

Content-Generation Engine
- Generate summarized versions of articles for different media
- Create SMS and Tweet with limited words and hashtags
- Create summarized paragraph per article to be sent as mail
- Connect with Translation Sub-engine for article language translation using Bing API

Distribution Engine
- Select articles per subscriber based on selected categories
- Push information to subscriber using preferred medium
- Post articles on twitter and facebook
- Generate and send digest e-mails based on subscriber preference

Conclusion and Future Work
KrishiEkta has been motivated by hope to provide the inaccessible information to users as even heading of a news item or an advisory method can make help a farmer in many ways. We’ve been running tests with some farmers and gardeners for the SMS model. E-mails are tested with faculty from universities and few members from the Agriculture Marketing Department. Our discussions with emerging farmer groups on twitter show positive signs of acceptance of this model.
KrishiEkta system needs to help the inner core of Indian farmers and therefore regional language support is needed in the system. Current version of translation sub-engine needs to be highly optimized for practical success. Also a lot more categorization will be created in the system allowing finer filtering of data. We will be increasing output media based on requirement and are trying to develop more encouraging ways of deployment to motivate users.

Highlights
- KrishiEkta gathers articles from multiple sources regularly in order to cover maximum information collection from web sources.
- The system categorizes news item in order to provide selective information to concerned subscribers.
- User interest in news and information rises more when specific categories of choice are available.
- Content-generation for multiple media tries to serve different user groups with ease of use towards respective medium
- For e.g., a group of farmers from Southern India active on twitter have expressed interest in news availability on farmer directed insurance schemes available in their area. Lack of information distribution has affected their needs in the past. With KrishiEkta, we can distribute such information to them through twitter and they can respond in time.
- Similarly when officials from Agriculture Marketing Department showed interest towards digest e-mails combining summarized news items in interest areas.
- Our attempt with translation and SMS service targets semi-literate farmers who can at least be informed about news related to their work and location.

KrisiEkta is a multiple level architecture, with four engines handling the main operations. First the Parsing Engine looks up through a list of sources and extracts all article content. Categorization Engine assigns relevant categories to every article. Content-generation Engine generates output strings for different formats and interacts with Translation sub-engine for language translation. Distribution Engine finally sends all information to subscribers in their desired format.