Building a Business Analytics Platform for Enhancing Commercial Beekeepers' Performance

Emergent Research Forum:

Decision Support and Analytics (SIGDSA)

Business Analytics for Managing Organizational Performance

Ethan Walton
Center for Analytics Research
Appalachian State University
Waltoneh@appstate.edu

Mark Jacobs
Center for Analytics Research
Appalachian State University
Jacobsms@appstate.edu

James Wilkes, Ph.D
Department of Computer Science
Appalachian State University
wilkesjt@appstate.edu

Joseph Cazier, Ph.D
Center for Analytics Research
Appalachian State University
cazierja@appstate.edu

Abstract

This research in progress paper plans to show how data analytics can be used to collect, manage and process large amounts of human interventive and natural order data, to build an intelligent hive management system to help optimize the performance of commercial beekeeping organizations to reduce bee colony losses, improve crop pollination services and improve sustainable food production.

Keywords

Data Platform, Analytics, Human Interventive Data, Natural Order Data, Smart Hives, Intelligent Hive Management System

Introduction

In agriculture, commercial and small farmers have only recently begun to recognize the benefits collecting and analyzing large amounts of data to optimize and expand production though the use of decision support and analytical technologies. Beekeepers, in particular, have not yet adopted the idea that data accumulation and analysis can create sustainable and profitable harvests. This is especially troubling given our dependency on bees for pollinating the bulk of our food supply and recent concerns about global bee health.

Industry Background

Apiculture, or beekeeping, is the study and management of bee colonies by humans. Unfortunately, the beekeeping industry has seen their subject matter slowly diminish over the years. The US has experienced a 40% decline in commercial honeybee population since 2006 and a 25% decline in similar populations since 1985 (Greenpeace, 2013). Though it is widely debated, many argue that the causes of the decline could be as diverse and widespread as disease, poor management, widespread chemical use, genetically modified crops, varroa mites, and changes in global temperature.

The impact of a declining bee population, as well as the other challenges that commercial beekeepers and growers face, is significant because pollinators have a profound influence on agriculture and the economy. In fact, one-third of all global agriculture relies on pollination from bees. Specifically, 87 of 115 global primary food crops require some type of pollination, with the largest percentage of dependency lying among nuts, vegetables, and fruits (Bauer and Wing, 2010).

The California almond industry, in particular, has significantly been impacted by the decline in bee health and bee population. Over 1,000,000 acres and 2 million bee colonies have been adversely affected, devastating California's largest export. Research indicates that in 2010, honeybees pollinated $12.4 billion worth of directly dependent crops and $6.8 billion worth of indirectly dependent crops in 2010 (Ramanujan, 2012). General equilibrium analysis estimates economy-wide impacts of bee losses at $334.1 billion when account for price effects on downstream sectors and households (Bauer and Wing, 2010).

The National Academy of Engineering examines related industries and concludes that there are several opportunities that could be taken advantage of through data collection, management, integration, and analysis. For example, precision agriculture uses tools and techniques, like big data and advanced analytics, to manage large farm areas in order to maximize food production, minimize environmental impact, and reduce costs. In recent years, commercial and small farmers have placed sensors, cameras, and computer screens in their tractors to optimize seed placement by GPS guidance. The sensors have the ability to monitor levels of certain elements such as nitrogen in the soil allowing farmers to take prescriptive action in response to field and soil metrics.

Precision agriculture applications have had compelling impacts on farms. For the 60% of farmers across the United States that have implemented precision agriculture, nationwide input costs decreased by 15 percent, crop yields increased by 13 percent, and water usage decreased by 50% (Bobkoff, 2015). In a case study quantifying savings for farmers in Australia, the Grains Research and Development Corporation found that on average, farmers saved $14-$30 per hectare using precision agriculture technologies. In addition, 60% of the farmers studied earned a profit and broke even from their initial capital investment in just under three years (Robertson, Carberry, & Brennan, 2007). The precision agriculture industry shows precedent that data analysis facilitates improved decision making, which leads to an increase in productivity.

Analytics Paradigm for Apiculture

The goal of analytics in apiculture is to go beyond reporting the current state of the hive to managing what the state of the hive should be given a certain set of conditions and inputs. This project seeks to analyze the benefits to building a unified data platform for the collection and analysis of data important to beekeepers and growers of pollinated crops. The data platform will be an integrated place to collect and store related data from various sources for the purpose of using analytics to better understand the data and then disseminating the knowledge gained to interested and relevant parties. This data would include human interventive inputs as well natural order data. Human interventive data is that data which is recorded as a result of the management of a hive. This includes inspections (queen status, frames of bees, food stores, presence of diseases or pests), hardware configurations, feedings, treatments, harvests, movements, and other management actions. In the future, once the platform is build and other parts of the project are finished, we may also be able to include external interventive factors such as nearby pesticide use, agricultural land use and history and other factors that may be relevant. Natural order data is data that is collected about things that happen in the natural world that affect beehives, pollinations or other related factors. This includes individual hive monitoring data like weight, temperature, humidity, sound, movement, and imaging as well as weather data.

Why Data Matters in Meeting these Objectives

The Analytics Pyramid (Figure 1) serves as a reference to the four types of analytics that will be used in this research. Our hope is to lay the groundwork to collect the required data in the descriptive phase that will allow us and the industry to move up the data analytics pyramid leading to smarter and better management practices optimized for economics and bee health and aided by smarter hives and hive support systems.

As we work with commercial beekeepers to measure the inputs and correlate them with hive outputs we will be able to use advanced data analytics techniques to better understand the economic and best practice management (BMP) tradeoffs to make better decisions for both bee health and beekeeper economics. As our breadth of beekeeping data begins to expand, predictive analytics will allow to predict US crop yields, measure environmental change and impact, and track and predict related economic trends. Just as with precision agriculture, prescriptive analytics would streamline the decision making process by prescribing the most optimal action to take given the decision alternatives. Herein, beekeepers would be able to automate complex decisions and tradeoffs to better manage their hives and thus improve the overall wellbeing of their bee population.

Figure 1: Analytics Pyramid

As we approach the peak of the analytics pyramid, will be able to go beyond reporting the current state of the hive to being able to manage the state of the hive given a certain set of conditions and inputs. This process would be enabled by the employment an intelligent software system that can help optimize global bee management. An intelligent hive management system is an advanced integrated software system that is necessary for the development and deployment of smart hives. Smart hives are beehives equipped with remote sensors that report information about the natural order data they are designed to collect. Smart hives go beyond collection and reporting of natural order conditions to actively monitoring those conditions, combined with human interventive data, to provide alerts and recommendations for improving the management of the hives or crops they pollinate. This is the brains of the organization that puts all of the various pieces together into a comprehensive system that will help the beekeeper optimize their operations for both economic and best management practices.

Process for Building the Analytics Systems

This study is designed as a natural longitudinal quasi-experiment that builds a data collection and management platform that will collect human interventive and natural order data from multiple commercial beekeepers that will be analyzed to improve our understanding of the economic and BMP tradeoffs beekeepers and farmers face. In the paragraphs below we will describe each step of the experiment to building this data platform, and how it will be used.

Step 1 - Data Standardization

The first step is data standardization. In this phase of the project we go through an open process of engaging key stakeholders, including researchers, the commercial beekeeping community, pollination dependent farmers and others. Additionally, this step will include scientific benchmarking to assist and validate that the right data is being collected in the right way. The goal of this process is to present and get feedback on the data that matters the most for the project and trying to define a process and procedure for what it means. This way there is a common and clear definition for all of the data going into the data platform.

Step 2 - Data Platform Modification

The second step is to build the data platform to collect, store and process the standardized data needed for this project in a way that will allow the application of advanced data analytics to generate key insights to build into improved software models. This step will include developing a prototype of the new system and seeking feedback from key stakeholders on how to modify and improve it.

Step 3 - Data Management and Integration

The third step is data management and integration. This includes bringing in all of the data into a unified platform that will store both human interventive and natural order data in a way that can be accessed and analyzed. This phase will also work to develop dashboards to display key performance information.

Step 4 - Advanced Analytical Techniques

The fourth step in the process is the application of advanced analytical techniques to the data collected and managed in the previous techniques. The primary goal of these analytical techniques is to gain knowledge and insights that can be integrated into an *intelligent hive management system* for commercial beekeepers that will help them balance economic and best management practices as well as help growers of pollination dependent crops.

Step 5 - Data Dissemination

The data dissemination step involves the sharing of knowledge and insights gained with key stakeholders. Steps four and five are linked in that as we learn things in step four we can disseminate or share them in step five. This will be a cycle in that once the data platform is built and the data integrated new insights and knowledge can be generated and shared with the community for the life of this project. The plan is to set up and participate in a web portal that will share this knowledge with the bee community.

The result of these 5 steps will be a comprehensive data platform that collects hive management and natural world data that can be managed, aggregated and analyzed to optimize commercial hive management.

Conclusion

The decline in bee health and population manifests a tremendous cause for concern in the agriculture industry. Pertinent action must be taken for the Common Good. We, as academics, have the power to help save the planet by enhancing the way we manage it. If our responsibility as scholars in the information system field lies in creating sustainable and environmentally sound development, then this project is an example of doing research in line with that responsibility. Beekeepers have a lot to gain from such efforts because the variety of data collection and analysis programs is currently limited, and our dependence on their work for food, make it a project with high impact potential.

REFERENCES

Bauer, Dana Marie and Ian Sue Wing. (2010). *Economic Consequences of Pollinator Declines: A Synthesis*. Web.

Bobkoff, Dan. (2015). *Seed by seed, acre by acre, big data is taking over the farm*. Web.

Greenpeace. (2013) "Bees in Decline: A review of factors that put pollinators and agriculture in Europe at risk", *Greenpeace Research Laboratories Technical Report (Review) 01/2013*, 2016. Web.

Ramanujan, Krishna. (2012). *Insect pollinators contribute $29 billion to U.S. farm income*. Web.

Robertson, M., Carberry, P., & Brennan, L. (2007, March 6). *The economic benefits of precision agriculture: case studies from Australian grain farms*. Web.

Shi, Yong. "Big Data History, Current Status, and Challenges Going Forward." The Bridge: Linking Engineering and Society 44.4 (2014). Web.

Watson, R. T., Boudreau, M., & Chen, A. J. (2010, March). *Information Systems and Environmentally Sustainable Development: Energy Informatics and New Directions for the IS Community*. Web.

Appendix A - Key Term Definitions

The following are some key terms and their definitions for the purposes of this paper.

Data Platform An integrated place to collect and store related data from various sources for the purpose of using analytics to better understand the data and then disseminating the knowledge gained to interested and relevant parties.

Human Interventive Data is that data which is recorded as a result of the management of a hive. This includes inspections (queen status, frames of bees, food stores, presence of diseases or pests), hardware configurations, feedings, treatments, harvests, movements, and other management actions. In the future, once the platform is build and other parts of the project are finished, we may also be able to include external interventive factors such as nearby pesticide use, agricultural land use and history and other factors that may be relevant.

Natural Order Data is that data that is collected about things that happen in the natural world that affect beehives, pollinations or other related factors. This includes individual hive monitoring data like weight, temperature, humidity, sound, movement, and imaging as well as weather data.

Smart and Intelligent Hives are beehives equipped with remote sensors that report information about the natural order data they are designed to collect. *Intelligent Hives* are hives that go beyond collection and reporting of natural order conditions to actively monitoring those conditions, combined with human interventive data, to provide alerts and recommendations for improving the management of the hives or crops they pollinate.

Analytics A formal methodology that uses a collection of computational processes and algorithms that follow a set of parameters defined and monitored by the user that are used to interpret, describe, discover, and identify patterns within and between structured and unstructured data.

An *Intelligent Management System* or *Intelligent Hive Management System* is an advanced integrated software systems that is necessary for the development and deployment of intelligent hives. This is the brains of the organization that puts all of the various pieces together into a comprehensive system that will help the beekeeper optimize their operations for both economic and best practices management.