Chapter XIII
Agricultural Data Mining in the 21st Century

E. Arlin Torbett
AgriWorld Exchange Inc., USA

Tanya M. Candia
Candia Communications LLC, USA

ABSTRACT

Data on the production, sale, repackaging, and transportation of fresh produce is scarce, yet with recent threats to national safety and security, forward and backward traceability of produce is mandatory. Recent advances in online marketing of fresh produce, a new international codification system and use of advanced technologies such as Radio Frequency Identification (RFID) and bar coding are working together to fill the gap, building a solid database of rich information that can be mined. While agricultural data mining holds much promise for farmers, with better indications of what and when to plant, and for buyers, giving them access to improved food quality and availability information, it is the world's health organizations and governments who stand to be the biggest beneficiaries. This chapter describes the current state of fresh produce data collection and access, new trends that fill important gaps, and emerging methods of mining fresh produce data for improved production, product safety and public health through traceability.

INTRODUCTION

Agriculture has always been, and continues to be, a key industry in the world's economy. In spite of its vital role in the world's economy, until very recently most data on production, brokerage, transportation, processing and consumption of agricultural products has been available only in the aggregate, rather than in a form that lends itself to investigation and learning. This is largely due to the fact that the majority of growers, shippers and buyers around the world are small businesses doing business by phone, fax or word of mouth.

Similarly, food traceability data related to product as it moves through the supply chain has been difficult to obtain with any certainly. Individual lots of fruits or vegetables are broken up, repackaged, and trans-shipped without any means of tracing forward or backward. This has profound implications for food safety and hygiene, and indeed to national defense.

Recent advances in online collection and maintenance of data related to fresh produce trading and exchange, as well as product inspection in the field and agricultural practices investigation, coupled with proposed standards for classification and Radio Frequency Identification (RFID) and other advanced tracking mechanisms have brought us to the point at which it becomes feasible to engage in effective data mining of agricultural data.

This ability has never been more important to a nation's and the world's interests. Such information can help to inform import and export policies; it can aid in efforts to ensure that agricultural workers receive fair pay for the fruits of their labor. Similarly, it can provide actionable information in the event of a food safety crisis. It can even assist governments to avoid or react to bioterrorism attacks. By correlating disparate data sources, agricultural data mining can finally lead to both forward and backward traceability, a thorny and hitherto unsolvable problem given the nature of the industry.

This chapter provides background and a review of the literature surrounding the topic, followed by an in-depth treatment of the recent historical context of agricultural data and problems associated with the lack of agricultural data repositories representing real-time data. It then discusses some important recent trends and issues that argue for development of online data bases related to agricultural production, movement and consumption. With a strong focus on the branch of agriculture related to fresh produce, the chapter discusses solutions and recommendations with regard to this highly demanding environment, where issues of determining optimal planting time, product and quantity relate to expected supply and demand; where product quality is inextricably tied to shelf life which drives pricing; where solid traceability is a key requirement for protection of the common good. Finally, the chapter presents future trends that will be important to research in the coming years.

BACKGROUND

As of 2006, an estimated 36 percent of the world's workers are employed in agriculture (International Labour Organization, 2007), down from 42% in 1996, making it by far the most common occupation across the globe. Although agricultural production accounts for less than five percent of the gross world product (an aggregate of all gross domestic products), it has been the world's key industry for centuries. It is the sale and purchase of agricultural products that has long provided the financial foundation for this industry.

Early History

Early agricultural transactions took place locally, between small growers and equally small buyers. Information about product quality, growing conditions, preparation and, especially, grower reputation was easy to come by, since eyewitness

Related Content