Data Integration and Management in Indian Poultry Sector

Susmitha Shankar and S. Thangam(✉)

Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Bengaluru, India
S_thangam@blr.amrita.edu

Abstract. Poultry sector- both commercial and smallholder- contribute significantly to both the financial stability of the poor and the economy of the nation. Record keeping in this sector not only ensures tight financial check of inventories but also helps in analyzing the factors affecting the quality and well-being of poultry. It is estimated that the world poultry production will increase many folds in the coming future, and so would the associated poultry health care analysis, inventory and infrastructure, climate analysis and other associated sectors. With the rise in population and growing demand for poultry product, manual record keeping and manual analysis of such huge, varied and dynamic data would turn out to be tedious and time-consuming task, causing out-of-time delivery of necessary decision. Thus, this sector is in need of software managed data analysis and data integration to ensure timely delivery of necessary action and most accurate prediction to help with the decision. We present a framework for data integration and analysis of poultry data to help in easy prediction, analysis and decision making in this sector, thus increasing the profits earned.

Keywords: Data integration · Poultry · Concatenation-based approach

1 Introduction

The transformation of Indian poultry from an unorganized and non-scientific sector to a commercial system by amalgamating scientific approaches has led to the growth in scientific discoveries in terms of poultry equipment and nutrients. Poultry data tracking can help with efficient tracking of the health of eggs and broilers. The advent of advanced and modern scientific equipment, methods and strategies in poultry industry brought about a lot of data processing and analysis for better poultry product production in terms of both quality and quantity. The global rise in egg production and consumption has been increasing. The enormous data would help in better decision-making faster and effectively.

Data analysis in poultry using computational techniques has shown remarkable results, indicating which factors largely affect the eggs and broilers and how much. Apart from the phenotypic and environmental factors that affect the poultry health, studies on the genetic structure of the poultry have also been made, thus helping in selection while breeding. The extensive use of mathematical and statistical approaches coupled with recent next-generation sequencing technology in the domain has helped in a detailed study of epidemiology [1]. Using the machine learning technique for

N. Kumar and R. Venkatesha Prasad (Eds.): UBICNET 2019, LNICST 276, pp. 96–102, 2019.
https://doi.org/10.1007/978-3-030-20615-4_7

automating the decision making on various aspects of poultry production cycle has been a keen research aspect for a decade. With highly efficient techniques such as SVM, regression models, HMMs, early prediction on the quality of the eggs and broilers could be made with higher accuracy and confidence [2].

With the advent of new technologies, the poultry data has grown enormous. Thus, the traditional manual method of record keeping is no longer suitable for decision making. Understanding the poultry to a maximum extent would require the comprehensive sight of all factors, including environmental factors such as the amount of ammonia in air, chlorine in water, number of flocks in unit kilometers as well as genetic factors. Although work on automated prediction of health and drops in egg production has been done, no work on the integration of poultry data has been found to the best of our knowledge. Processed data from multiple sources integrated into one platform could significantly increase the predictive analysis. An efficient platform for dynamic integration of the poultry data coupled with the visualization tool could help the poultry farmers to yield better decisions. Thus, this work aims at the creation of a framework for integration and visualization of poultry data.

The understanding of the domain for the work is taken from various articles and published materials as presented in Sect. 2. Followed by Sect. 2, the proposed architecture is presented in Sect. 3 and Sect. 4 deals with the implementation. Section 5 concludes the work.

2 State of Art

The appropriate synchronization of the two sectors, namely the commercial sector and the unorganized sector, helps in enhancing the economic conditions of the smallholders. The United State Department of Agriculture (USDA) Global Agricultural Information Network (GAIN) report for 2015 and 2016 estimates per capita consumption of poultry meat to be around 3.1 kg per year and 3.6 kg per year respectively and was forecasted to rise by 7% over 2016. Also, the per capita consumption of eggs in India is reported to be 62 eggs per year. However, both consumption of meat and egg is considerably lower compared to the world average (17 kg per year for meat), which establishes the possibility of the further flourishing market [3]. With the rise in income and demand, the need for healthier products would also increase. Active research and development facility from governmental and private sectors to develop quality and quantity of poultry product would further raise the poultry sector.

The profitability in the poultry sector also depends on the ease of vertical and horizontal integration [4, 5] and the support of governmental agencies for the same. Vertical integration in the poultry sector refers to one company many controls policy, wherein one company owns multiple stages of productions such as feed mills, hatcheries and many more. This leads to independent complete production groups, the better utility of space, enhanced quality control, and comparatively higher productivity and profitability. These are also referred to as “integrators”. Horizontal integration helps in rearing different types of dependent livestock rearing such as poultry with fisheries, etc. The interdependencies of various livestock enhance the symbiotic existence, quality and the by-product such as generated wastes, have higher nutritive proportion and are

best in quality. Although there is a high scope of profitability in the poultry sector in India, there are also challenges that have to be seriously dealt with.

Diagnosis of the health of poultry at the appropriate time, efficient management and reuse of bio-waste, proper storage, and transport facilities, availability of quality supplements and many more factors. To cope with the challenges listed above, the Government of India has formulated infrastructural regulations as mentioned in Poultry Farm Manual (2014–15) by Department of Animal Husbandry, Dairying and Fisheries, Ministry of Agriculture and Farmers Welfare, Government of India, which includes compliance of location of farm with the State Pollution Control Board, minimum geographical distance of more than 100 m from major water drinking source and 500 m from other poultry enterprise source, proper drainage outlet, dimension of the farm, etc. Strict check and regular follow-ups on these rules must be initiated [6]. Also, there are few international agencies ready to join hands with Indian Poultry sector to yield necessary technical guidance to enhance this field [7].

However, voluntary check by the organizations and small-scale farmers along with motivation by the governmental and non-governmental organization would prove to be much beneficial and effective. The analysis and decision making in this sector involve a lot of factors to be considered. Missing a few factors would reduce the efficiency of the farm product, leave a bad influence on the farm and poultry product and may also affect the health of the poultry. Thus, a software managed check and automated decision analyzer could prove to be a great help.

Various case studies in automated livestock handling have been done, all of which points towards the need for software enhanced study [8]. Studies on the economic condition bought due to the rearing of poultry, a harmful environmental condition due to the emission of ammonia from the pig farm, efficient routing technique for transportation, maintaining inventory management according to the supply-demand chain are studied in detail [9, 10]. Most of the decision made in traditional livestock rearing set-up are sequential and not strategic. The strategic decision could yield more profit. Stable strategies, one which aims at the sustainability of the production system while maximizing the profit, are also emerging research goal [11]. Many integrative methods are used to create multi-model strategic plans including the use of fuzzy logic.

Programming models target efficient supply-chain decision making in seasonal and non-seasonal supply chain products [12]. The use of a Bayesian network to analyze the supply chain demand, thereby deciding whether to enter the market on the basis of available resources has been studied by Laper et al. [13]. Machine learning approaches such as SVM, are used to analyze the factors while making the decision on commercial egg production and broiler management [14, 15].

Mathematical and statistical approaches are discussed in detail to analyze various aspects of egg production, broiler feeding, accurate cut while extracting meat, inventory management and other aspects of poultry rearing [16]. Many artificial intelligence tools can also be made use of to enhance high-quality poultry product.

3 Proposed System

The framework aims at creating a data integration platform for unified and complete analysis of poultry data to extract valuable inferences. It is composed of 2 layers: the presentation layer and the processing layer. The presentation layer deals with taking inputs from the consumers, who could be poultry farmers, researchers, government agencies, etc., form queries to extract result and also to display those results. This could also be used by the producers of the information to put the data into the processing unit. The processing unit deals with the data integration platform. The data integration platform extracts the time-series and high-dimensional poultry information from the producer, make quality check on the basis of the error rate for the data, prepares the data by cleaning and processing for redundant and inconsistent data, choosing appropriate data integration technique on the basis of the data to be integrated, applying the integration technique and finally checking the quality of the integrated system.

The consumer enters the keyword related to which the inference is to be extracted. These keywords are converted to query, which is then run under the corresponding database to extract information. The integrated database would be formed when the producers or suppliers of the data gives the data to the platform which goes through correction, preparation and integration process, and is stored in the database. Figure 1 shows the diagrammatic representation of the proposed architecture.

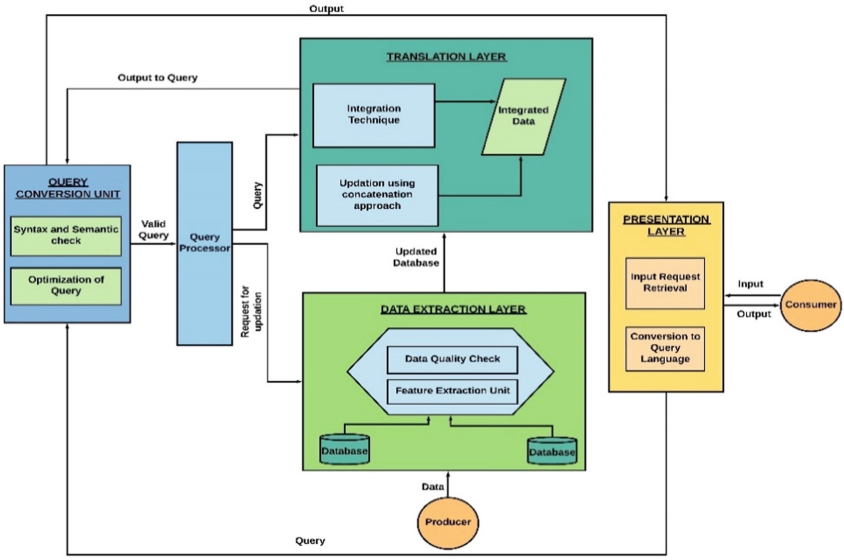

Fig. 1. Diagrammatic illustration of the proposed architecture.

4 Implementation

The data were extracted one by one from the data repository in raw format, mostly in json and xlsx. The data sources used in the integration was the open source poultry data from data.gov.in [17] and indiastats.com [18]. The first step involved the analysis of the data and the interpretation of different features associated. The statistical analysis such as correlation among the data, was conducted using R. The data represented longitudinal aspects, that is country wise or region wise studies, statistics of backyard desi and improved poultry for 2016–17, statistics of the actual the estimated production of poultry from 1993 to 2014, per-capita egg production from 1950 to 2018 for the state of Tamil Nadu and the total production of eggs for all Indian States. The challenge of missing values, due to unavailability of information for various time range, were fixed by referring to multiple sources for creating a consistent model. Once the data was extracted from their respective sources, the first step was to standardize the data.

These data were then segregated into different classes based on their interpretation, significance and the relation between them to multiple classes, programmatically using C# in .Net Core environment. Since the data separated into logical chunk involved a large amount of data operations, the code was run in an asynchronous parallel fashion. The segregation was done in such a way that the inconsistency is reduced. This was achieved by normalizing the data to 3NF. Concatenation based approach was used as the integration technique for this model. The data from multiple sources in varied format were concatenated into a logical unit using this approach. Once the classes were created and the data was divided, a SQL connection was established with SQL Server, where the tables reside, using Microsoft Azure Data Studio. Figure 2 shows the entity class diagram for the integrated poultry. The framework developed is feasible and highly scalable.

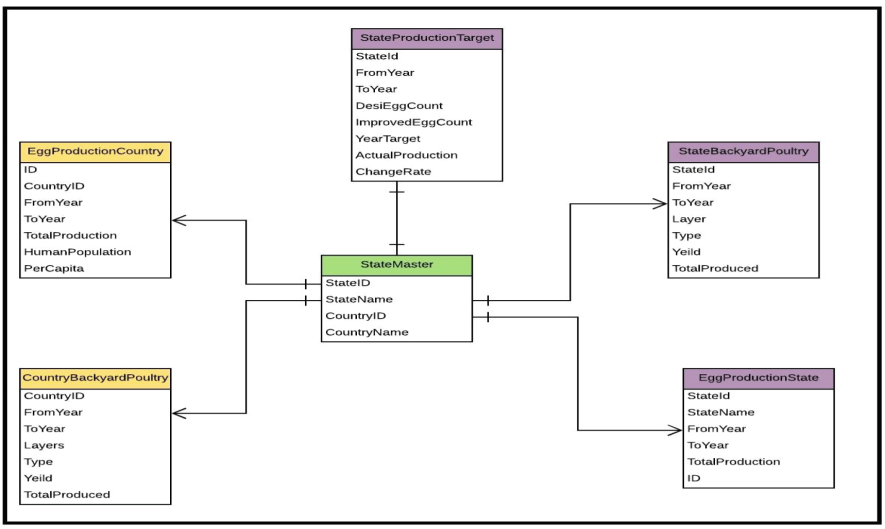

Fig. 2. Entity-class diagram for integrated poultry database

5 Conclusion and Future Work

The Poultry farming is one of the important means by which the rural community can be uplifted. The data from poultry is rich, dynamic, large and from varied sources. For proper analysis, the integration of this data is the need of the hour. Effective integration could be made after ensuring the quality of each of the source of the data and pre-processing. Automation of the task of integration would be helpful in making logical and efficient decision to the utmost level of confidence, even in the absence of a human. The choice of implementation of a machine learning algorithm for the integration of this multi-sourced varied data would show a promising result.

The proposed architecture deals with the creation of an integrated system for poultry data from multiple sources coupled with the visualization tool. The system so integrated is present in the Relational Database Management System. Although the system is implemented for the data of Tamil Nadu, the study can be extended to multiple regions, providing better longitudinal studies and analysis. This data could be made public for the researchers to analyze the state of art for poultry for a particular region. The creation of better model requires the availability of open source data in poultry. The availability of a larger amount of data could pose a big-data challenge. This would create a strong load on the system. Thus, the future scope involved in the project is the use of distributed big-data technology such as Hadoop Distributed System, to ensure efficient model even if the data in the integrated system explodes.

References

1. Kao, R.R., Haydon, D.T., Lycett, S.J., Murcia, P.R.: Surprise me: how whole-genome sequencing and big data are transforming epidemology. Trends Microbiol. **22**, 282–291 (2014)
2. Hepworth, P.J., Nefedov, A.V., Muchnik, I.B., Morgan, K.L.: Broiler chickens can benefit from machine learning: support vector machine analysis of observational epidemiological data. J. R. Soc. Interface **9**, 1934–1942 (2012)
3. Intodia, V.: Poultry And Poultry Products Annual 2016, Usda Gain Report, New Delhi (2016)
4. Olasunkanmi, M.B., Dayo, O.A., Philips, S.M.: Vertical integration and technical efficiency in poultry (egg) industry in Ogun and Oyo states, Nigeria. Int. J. Poult. Sci. **5**, 1164–1171 (2006)
5. Moses, B.O., Olanrewaju, O.A., Olubanjo, I.A.: Economics of horizontal integration in poultry industry in South-West Nigeria. Int. J. Poult. Sci. **11**, 39–46 (2012)
6. Viaene, J., Verbeke, W.: Traceability as a key instrument towards supply chain and quality management in the belgian poultry meat chain. Supply Chain. Manag. Int. J. **3**, 139–141 (1998)
7. Our Bureau. Good Scope for FDI in Poultry Sector, Says Assocham @Businessline (2018). https://www.Thehindubusinessline.Com/Economy/Agri-business/Good-scope-for-fdi-in-poultry-sector-says-assocham/Article20348014.Ece1. Accessed 25 Aug 2018
8. Rebolledo, B., Gil, A., Pallarés, J.: A spatial ammonia emission inventory for pig farming. Atmos. Environ. **64**, 125–131 (2013)

9. Song, Q., Gao, X., Santos, E.T.: A food chain algorithm for capacitated vehicle routing problem with recycling in reverse logistics. Int. J. Bifurcat. Chaos **25**, 1540031 (2015)
10. Ten Napel, J., Van Der Veen, A.A., Oosting, S.J., Groot, P.W.: A conceptual approach to design livestock production systems for robustness to enhance sustainability. Livest. Sci. **139**, 150–160 (2011)
11. Teimoury, E., Jabbarzadeh, A., Babaei, M.: Integrating strategic and tactical decisions in livestock supply chain using bi-level programming, case study: iran poultry supply chain. PLoS ONE **12**, E0185743 (2017)
12. Baky, I.A.: Solving multi-level multi-objective linear programming problems through fuzzy goal programming approach. Appl. Math. Model. **34**, 2377–2387 (2010)
13. Sauian, M.S., Othman, N.C.: Optimizing livestock rearing using MCDM. In: 2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE) (2012). https://doi.org/10.1109/icssbe.2012.6396584
14. Morales, I.R., Cebrián, D.R., Blanco, E.F., Sierra, A.P.: Early warning in egg production curves from commercial hens: a SVM approach. Comput. Electron. Agric. **121**, 169–179 (2016)
15. Fialho, F.B., Ledur, M.C.: Segmented polynomial model for estimation of egg production curves in laying hens. Br. Poult. Sci. **38**, 66–73 (1997)
16. Barge, P., Gay, P., Merlino, V., Tortia, C.: Radio frequency identification technologies for livestock management and meat supply chain traceability. Can. J. Anim. Sci. **93**, 23–33 (2013)
17. data.gov.in. (Internet). https://data.gov.in/. Accessed 14 Jan 2019
18. Socio Economic Statistical Data & Facts About India - e-Resource Library for Research | Indiastat (Internet). https://www.indiastat.com/. Accessed 14 Jan 2019