S02(T)-OP-4

big Data analytics in the animal production domain

Claudia Kamphuis, Erwin Mollenhorst, Roel Veerkamp
Animal Breeding and Genomics, Wageningen University & Research, Wageningen, Netherlands

The term Big Data was introduced to the high-tech community in the early 1990s, and was commonly described by three main-stream characteristics (Volume, Variety, and Velocity). Advances in computer storage and processing power, the sharp reduction in costs of sensors and communication tools, and developments in the Internet of Things all contributed to an exponential use of the term Big Data. Probably associated as being in the 'peak of inflated expectations' by the Gartner Hype Cycle. Big Data, and its associated analytics, has also been an increasingly used within the animal production domain, where it is described as highly valuable. The on-going technological developments in the animal production domain, indeed cause an increase in sources that generate data. These sources include machines (e.g., milking robot), fields (e.g., soil type and temperature), animals (genomic data, sensors), environment (e.g., weather), and the production chain (e.g., slaughter information). This availability of increasing volumes of high-velocity data from a variety of data sources is unique compared to most fields outside agriculture and attract the attention of Big Data analytical skills. We participate in projects combining genomic and sensor data to understand and predict complex traits like longevity, resilience, and efficiency in dairy cattle. Field and environmental data to improve fertilization of fields, adding to a circular agriculture. Animal data and data from the production chain are used to develop decision support systems in the pig industry (monitoring health at national level, or that aid farmers in grouping piglets for a more uniform batch). During these projects we experienced key factors to successful benefit form Big Data analytics in the animal production domain. Firstly, data have to be made available for the benefit of farmers, consultants, and legislation. Secondly, Big Data analytics is more than just giving the data to data scientists for analyses. It requires a multidisciplinary team of IT specialists that deal with storage and access to these large volumes of data, data scientist to prepare and analyse it, and domain knowledge to optimise the data preparation and to interpret the results. Lastly, Big Data analytics may require, or at least benefit, from other working routines. Examples of these new working routines include hackathons, a pressure cooking setting where multidisciplinary teams have to combine domain knowledge, data, and hard- and software tools during 24-36h competition. In conclusion, due to the on-going technological developments in the animal production domain, more Big Data will come available. But technology is not the silver bullet for success. To become successful, however, sharing data is essential, as well as a good understanding and collaboration between IT, data science, and domain knowledge.

Keywords: Big Data, data analytics