



Conference Proceedings | DOI: <https://doi.org/10.18174/FAIRdata2018.16261>

Developing new data driven models using Bayesian Network approach and machine learning methodologies

Marvin, H.¹ & Y. Bouzembrak¹

¹ RIKILT, Akkermaalsbos 2, 6708 WB Wageningen, the Netherlands
Corresponding author's e-mail: hans.marvin@wur.nl

The food supply chains are complex and many drivers such as climate, human behaviour, economy have a direct and indirect effect on the performance of the supply chains. It has become clear that a system approach is needed to solve the problems related to the production of healthy and nutritious food and to tackle food safety problems. In such system approach, it is clear that big data technologies and tools are needed, including a safe and powerful infrastructure to allow the handling of big data and to ensure interoperability. This project will explore the use of machine learning methodologies to predict food safety in particularly exploring the possibilities of Bayesian Networks (BN). For this purpose, we selected food safety in feed of dairy cows in the Netherlands and the impact of climate and use of agrichemicals as a case. Data was collected for the years 2000-2013 and used to construct a BN, which was optimised for the prediction of the contamination level (i.e. presence of contamination). The overall accuracy of prediction of the level of contamination in feed was 90.3% and sensitivity analysis could show the impact on the contamination of the individual input parameters (e.g. climate and chemical use).