Chapter 7
Machine Learning Models for Yield Prediction Based on Environmental Data

Volodymyr Kravchenko
https://orcid.org/0000-0002-8033-3985
National University of Life and Environmental Sciences of Ukraine, Ukraine

Roman Rudenskyi
https://orcid.org/0009-0002-3682-9702
National University of Life and Environmental Sciences of Ukraine, Ukraine

Semen Voloshyn
https://orcid.org/0000-0002-4913-7003
National University of Life and Environmental Sciences of Ukraine, Ukraine

Valentyna Korolchuk
https://orcid.org/0000-0002-3145-8802
National University of Life and Environmental Sciences of Ukraine, Ukraine

Tetiana Voloshyna
https://orcid.org/0000-0001-6020-5233
National University of Life and Environmental Sciences of Ukraine, Ukraine

ABSTRACT

This chapter explores the use of machine learning for predicting crop yields based on environmental data. Key factors contributing to its potential include the availability of large volumes of historical data and powerful algorithms capable of processing it. Changes in climatic zones make artificial intelligence crucial for enhancing productivity. Procedures for data preparation are proposed, considering

DOI: 10.4018/979-8-3373-0842-5.ch007

various factors such as climate conditions and fertilizer usage, improving predictions' accuracy. The systematization of these procedures creates a unified database for the agricultural sector, optimizing cultivation and minimizing risks. Advanced approaches and algorithms and the opportunities and challenges of integrating them into modern agrarian technologies are examined. A cloud-based solution architecture is proposed, ensuring flexibility in analytics. This will benefit farmers, researchers, and investors in enhancing resilience and productivity.

INTRODUCTION

Challenges in the agricultural sector, including climate change and limited resources, require implementing modern innovative solutions, including those based on artificial intelligence (AI), which will help increase agricultural production's productivity and efficiency. In the context of the rapid development of AI technologies and their impact on farming practices, exploring the possibilities of applying such technologies in the agricultural sector is a highly relevant and vital issue. Agriculture faces numerous challenges, such as climate change, declining soil fertility, and resource scarcity, which require new approaches to ensure sustainability and productivity. Yield forecasting is a key element that helps farmers and agribusinesses optimise and improve their performance, reduce risks, and make informed decisions about crop planning, management of agronomic measures and techniques, etc. Machine learning models offer potent tools for yield forecasting, allowing for complex interactions between different factors and adaptation to changing environmental conditions.

BACKGROUND

Modern machine learning methods are being widely used in agriculture and agricultural production, in particular, to predict crop yields, diagnose and predict possible plant diseases, analyse and manage the efficiency of production processes in various agricultural sectors, etc.

Machine learning algorithms, such as k-nearest neighbours, random forest, support vector machine, and gradient boosting, allow for effective prediction of crop yields with MAPE (Mean Absolute Percentage Error) up to 10% for making management decisions in the agricultural sector (Sabitov et al., 2023). Machine learning techniques can improve the efficiency of smart farms by managing heterogeneous information and data, increasing profitability, reducing costs, predicting crop yields, and transforming sensor data (Balducci et al., 2018). Using data mining and