

ПРОГНОЗИРОВАНИЕ ПРОИЗВОДСТВА СЕЛЬСКОХОЗЯЙСТВЕННЫХ КУЛЬТУР С ИСПОЛЬЗОВАНИЕМ МАШИННОГО ОБУЧЕНИЯ

Сильницкий И.С. Университет ИТМО

Научный руководитель – доцент, Маргун А.А.

Университет ИТМО

Прогнозирование является одним из основных этапов управления предприятием, если нельзя спрогнозировать ход развития события, предсказать развитие какой-либо ситуации, достаточно сложно ставить четкие цели, принимать управленческие решения и оптимизировать работу людей. Прогнозирование в сельскохозяйственной отрасли занимает важную роль, так как продукция имеет срок годности и ее переизбыток может привести к большим потерям, поэтому данный вопрос очень актуален на данный момент.

В работе была найдена база данных из открытых источников по производству молока на сельскохозяйственном предприятии. Разработаны LSTM и SARIMA модели прогнозирования с использованием машинного обучения. Далее произведено компьютерное моделирование с целью получения графических и численных результатов, полученные модели показали достаточно близкие результаты.

Введение. Сейчас все больше компаний старается строить прогнозы, что бы повысить свою прибыль, поэтому предложенных способов достигать точных результатов все больше, данная область активно развивается и становится неотъемлемой частью процесса планирования. Были изучены предложенные зарубежными и отечественными исследователями варианты решения данной задачи в различных областях и принято решение разработать модель прогнозирования для сельскохозяйственной отрасли. После изучения методов решения задач прогнозирования были выбраны LSTM и SARIMA модели. Выбор в пользу этих моделей был обусловлен возможностью работы с большими временными рядами и взаимосвязью всех предыдущих значений с прогнозируемым. Данные для исследования были найдены в открытых источниках.

Основная часть. В основе работы с данными лежит анализ полученной базы. Была проведена предобработка данных, в которой датасет был изучен на аномальные значения, проверен на наличие дубликатов и сменены типы данных где это требовалось. Затем база данных была проиндексирована по временным отметкам и задана частота попадания данных в таблицу.

Основным этапом разработки моделей является подбор параметров. В модели SARIMA подбор параметров реализуется программно с помощью информационного критерия Акаике. В модели LSTM выбор параметров производился путем подбора, что бы полученная функция была максимально приближена к функции правдоподобия.

Затем были построены модели прогнозирования. Для определения точности работы моделей все данные были разбиты на две части. Первая часть данных отвечала за обучение и именно по ним происходило обучение моделей, а вторая часть данных послужила эталонным результатом, с которым сравнивался прогноз каждой модели для выявления точности полученных результатов. Оба прогноза были на интервал в один год. Было сделано компьютерное моделирование и полученных численные результаты, построены графики. Далее модели сравнивались с эталонными значениями и между собой. Обе модели показали близкие результаты, погрешность у SARIMA и LSTM моделей составила не больше 8% от реальных значений.

Выводы. Работа была реализована на реальных данных, полученных из открытых источников, погрешность результатов составила не более 8%. Исследование проводилось в области сельскохозяйственного сектора и пригодно для внедрения на реальное предприятие.

Сильницкий И.С. (автор)

Подпись

Маргун А.А. (научный руководитель)

Подпись