Using Deep Learning to Predict Plant Growth and Yield in Greenhouse Environments

Bashar Alhnaity [a], Simon Pearson [b], Georgios Leontidis [a] and Stefanos Kollias [a]
School of Computer Science, University of Lincoln, Lincoln, LN6 7TS, UK.[a]
Lincoln Institute for Agri-Food Technology, University of Lincoln, Lincoln, LN6 7TS, UK.[b]

Abstract

Effective plant growth and yield prediction is an essential task for greenhouse growers and for agriculture in general. Developing models which can effectively model growth and yield can help growers improve the environmental control for better production, match supply and market demand and lower costs. The proposed study utilises ML and DL techniques to predict yield and plant growth variation across two different scenarios, tomato yield forecasting and Ficus benjamina stem growth, in controlled greenhouse environments. Very promising results, based on data that have been obtained from two greenhouses, in Belgium and the UK are presented.

Introduction

As with many bio-systems, plant growth is a highly complex and dynamic environmentally linked system. Therefore, growth and yield modeling is a significant scientific challenge . Modeling approaches vary in a number of aspects (including, scale of interest, level of description, integration of environmental stress, etc.). According to [1], there is a large number of tools that can help farmers in making decisions. These can provide yield rate prediction, suggest climate control strategies, synchronise crop production with market demands. A deep learning model is proposed in this paper, which is trained with environmental (CO_2, humidity, radiation, outside temperature, inside temperature), as well as, actual yield and stem diameter variation measurements and has the ability to produce accurate prediction of either ficus stem diameter, or tomato yield problems.

Models

SVR: Support Vector Regression projects the input data into a higher dimensional space using kernel function and separates different classes of data using a hyperplane.

RF: Random Forest uses decision trees. The idea of ensemble learning is that a single predictor is not sufficient for predicting the desired value of test data.

LSTM: Long Short Term Memories model was initial introduced in [2] with the objective of modeling long term dependencies and determining the optimal time lag for time series problems.

References

[1] Qaddoum, Kefaya and Hines, EL and Iliescu, DD *Yield prediction for tomato greenhouse using EFuNN*, ISRN Artificial Intelligence (2013)

[2] Hochreiter, Sepp and Schmidhuber, Jürgen *Long short-term memory*, Neural computation (1997)

Acknowledgements

This work is part of EU Interreg SMARTGREEN project (2017-2021). We would like to thank all the growers (UK and EU), for providing the data; Their valuable feedback, suggestions and comments are highly appreciated to increase the overall quality of this work.

Results and Discussion

We have developed and tested DL (LSTM), SVR and RFR prediction models to predict plant yield and growth in greenhouse environments for:

a) ficus growth prediction based on the SDV indicator:

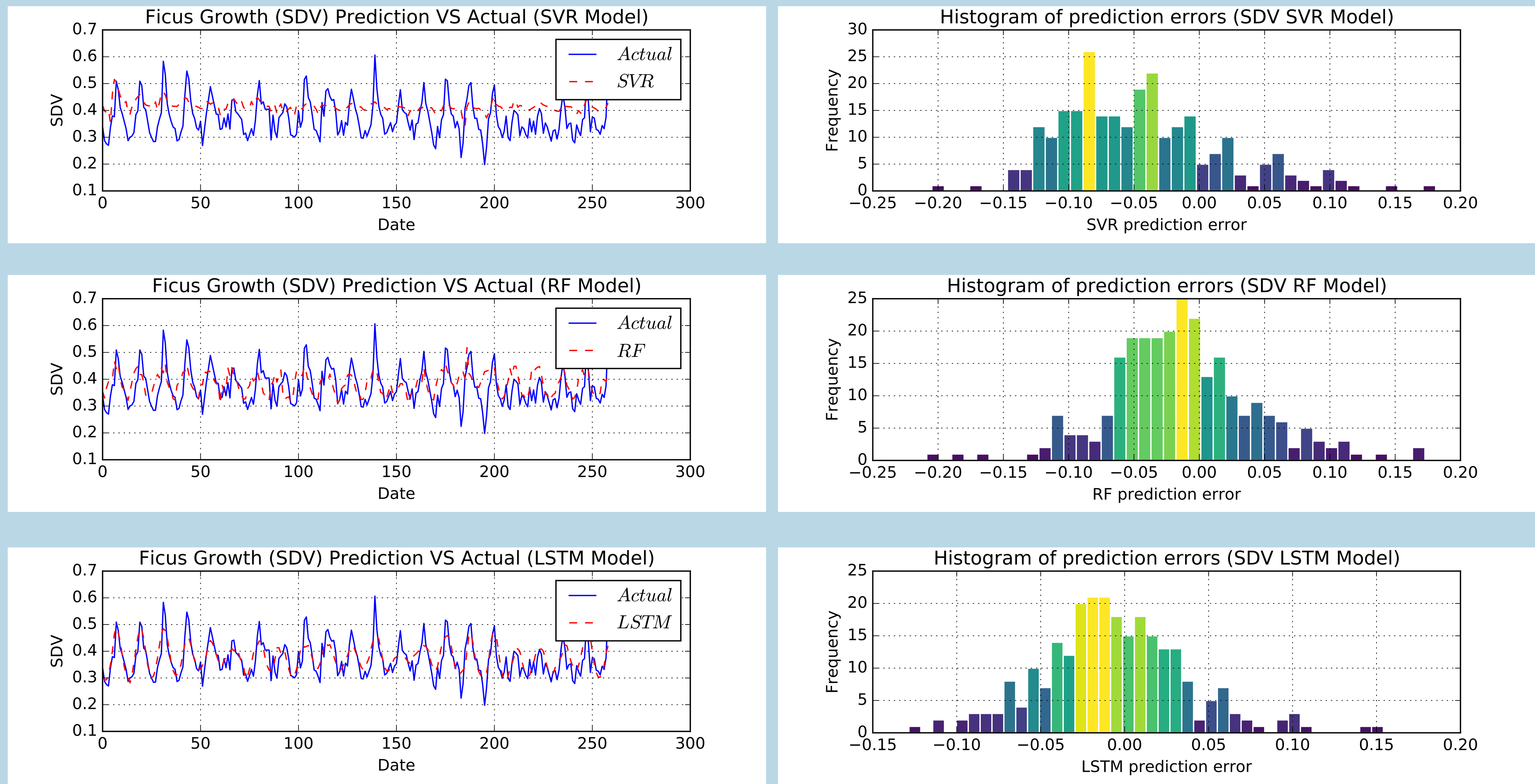

Figure 1: Testing results and performance comparison of Ficus growth (SDV) predictions.

b) tomato yield prediction:

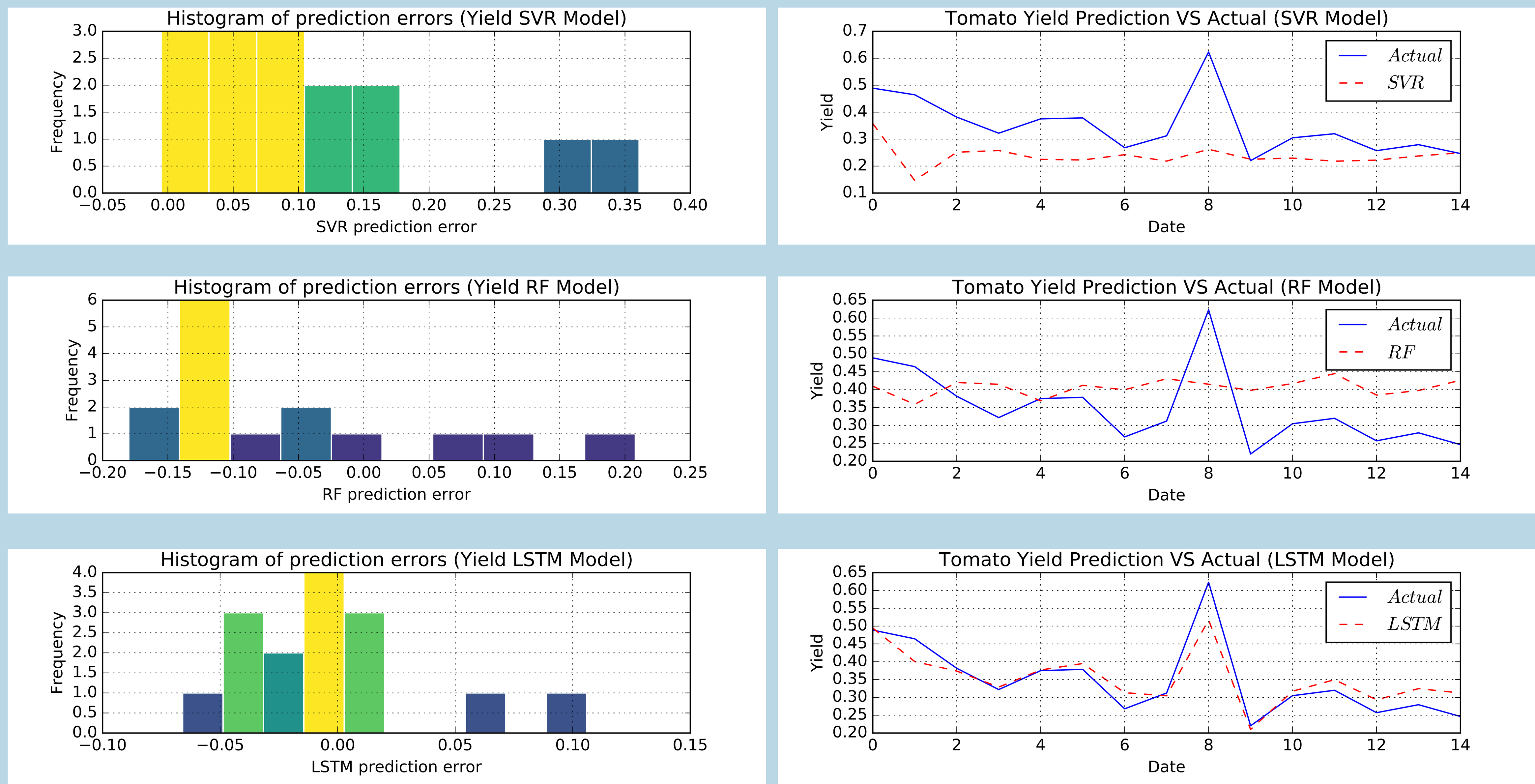

Figure 2: Testing results and performance comparison of Tomato Yield predictions.

The obtained results clearly show that the DL/LSTM model outperforms the SVR and RF ones, in both experiments. Table 1 shows the obtained accuracy, in terms of MSE, RMSE and MAE, when each of the (trained) three models is applied to the test datasets, in both experiments.

Datasets	Tomato Yield			Ficus Growth(SDV)		
Models	SVR	RF	LSTM	SVR	RF	LSTM
MSE	0.015	0.040	0.002	0.006	0.006	0.001
RMSE	0.125	0.200	0.047	0.073	0.062	0.042
MAE	0.087	0.192	0.03	0.070	0.063	0.030

Table 1: Performance of the DL/LSTM model for plant yield and growth prediction

Conclusions

The paper developed a DL approach using LSTM for Ficus growth (represented by the SDV) and tomato yield prediction, achieving high prediction accuracy in both problems. Experimental results were presented that show that the DL technique (using a LSTM model) outperformed other traditional ML techniques, such as SVR and RF, in terms of MSE, RMSE and MAE error criteria. Hence, the main aim of our project is to develop DL methodologies to predict plants growth and yield in greenhouse environment. Future studies looking at the continuity of : a) greatly increase the number of collected data that are used for training the proposed DL methods; b) extending the DL method so as to perform multi-step (at a weekly, or a multiple of weeks basis) prediction of growth and yield in a large variety of greenhouse, in the UK and Europe.