Prediction of frost events using machine learning and IoT

Ana Laura Diedrichs[1] Facundo Bromberg[1,2] Diego Dujovne [3]

[1]Universidad Tecnológica Nacional – Facultad Regional Mendoza , Argentina - [2]CONICET, Argentina - [3]Universidad Diego Portales, Chile

What's the problem with frost events?

In a couple of hours, farmers can lose everything because of frost events. Plants and fruits suffer from frost events as a consequence of water icing inside the internal tissues present in the trunk, branches, leaves, flowers and fruits. However, water content and distribution is different among them, generating different damage levels. Handling frost events is possible by using many countermeasures such as heating or removing the surrounding air among crops. Given the socio-economical implications of this problem, there have been some efforts to design a system to predict frost events, but with partial success and the prediction performance is poor ($R^2 \approx 0.6$).

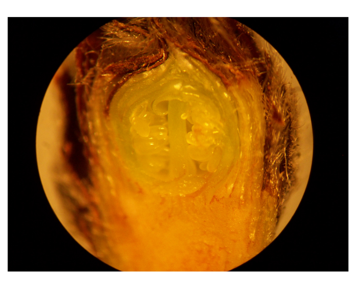

Figure 1:Flower bud on vegetative rest

Figure 2:Flower bud prior to spring blooming

Figure 3:Peach flower buds at different stages of winter rest. The flower bud has a lower resistance to frost in (b).

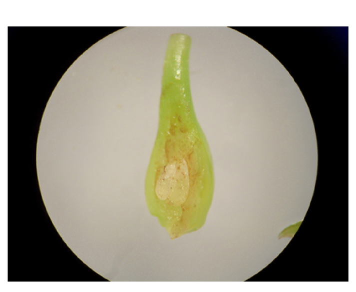

Figure 4:healthy.

Figure 5:damaged

Figure 6:Gynoecium healthy and damaged by frosts. Damage is visualized by brown tissues.

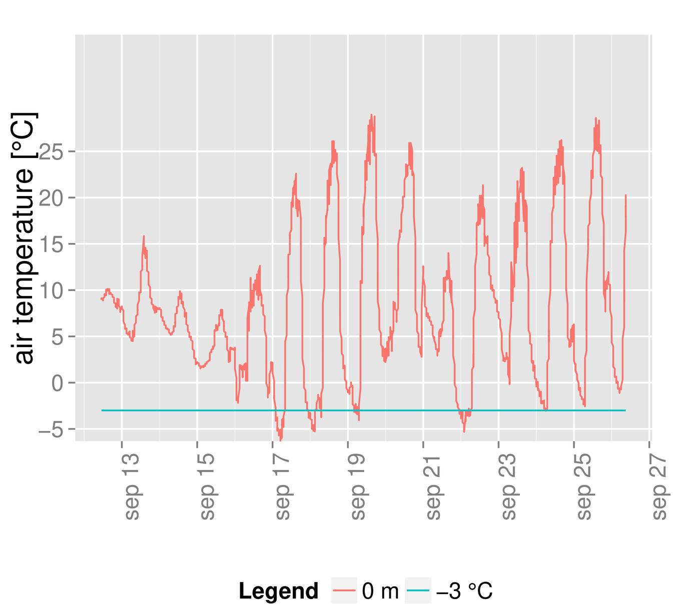

Figure 7:Temperature trend during a week of frost events in September 2013, Junin, Mendoza

Dataset

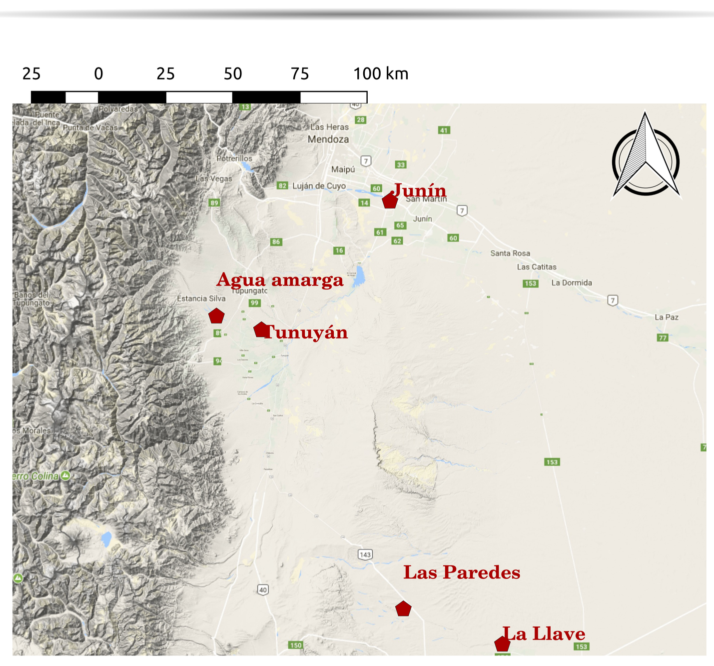

Figure 8:Map of the DACC's meteorological stations located in Mendoza, Argentina.

Download paper

Diedrichs, A. L., Bromberg, F., Dujovne, D., Brun-Laguna, K., Watteyne, T. (2018). Prediction of frost events using machine learning and IoT sensing devices. IEEE Internet of Things Journal, 5(6), 4589-4597.

We use machine learning for:

- To predict if it is going to frost tomorrow in a location given the temperature and the humidity from the station and/or the other surrounding stations. (classification)
- To predict the minimum temperature for the next day. (regression)

Datasets considerations

- We have to deal with an unbalanced dataset. So, we propose that using SMOTE (Synthetic Minority Oversampling Technique)could improve the sensitivity of the model.
- We want to see if the information of the neighbor stations helps to improve the prediction or not. (if they are essential features).
- We want to know if only the temperature information are enough to predict (cheap sensors).
- To decide if we need information from all the seasons or just from the Spring season.

Experiments

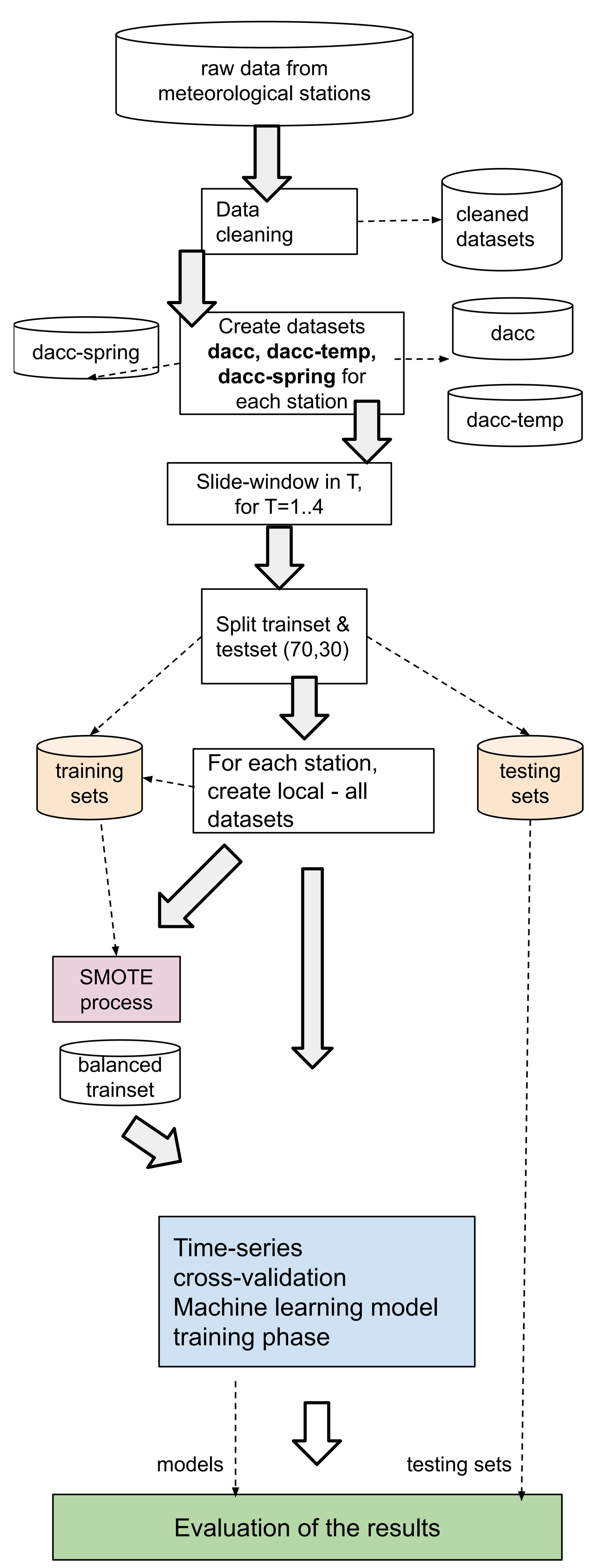

Figure 9:Experiment work-flow diagram.

Thanks to

Results

We evaluate the regression results using the following metrics:

- Root Mean Square Error $RMSE = \sqrt{\frac{1}{n}\Sigma(Y_{pred} - Y_{real})^2}$
- Mean Absolute Error $MAE = \frac{1}{n}\Sigma_{t=1}^{n}|Y_{pred} - Y_{real}|$
- Sensitivity: $\frac{TP}{TP+FN}$ also known as true positive rate, probability of detection and recall. Higher values of sensitivity indicates that we have a good predictor of the positive class.
- Precision: $\frac{TP}{TP+FP}$ reflects how accurately is the predictor for predicting the positive class. The higher is this value the lower the chances of false positives.
- F1-score: $F1 = \frac{2 \cdot precision \cdot sensitivity}{precision + sensitivity}$, which represents a balance between precision and sensitivity.

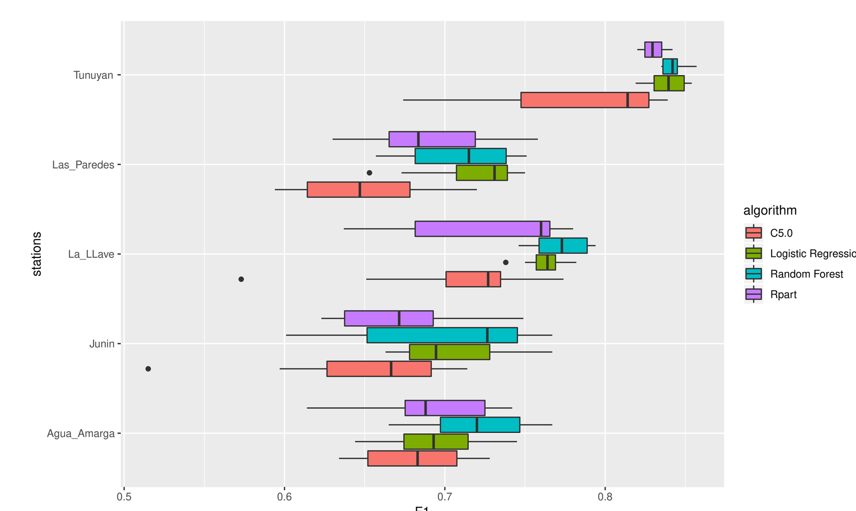

Figure 10:Classification models performance

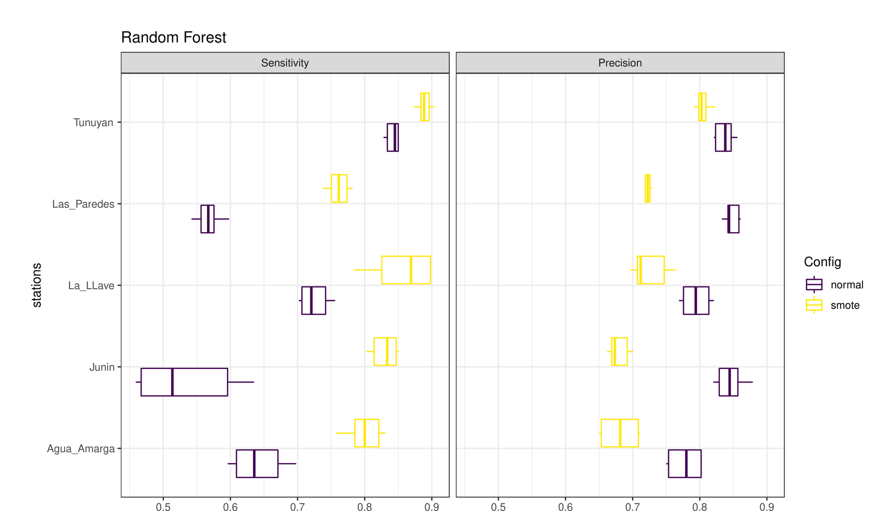

Figure 11:Random forest classification: SMOTE vs normal training

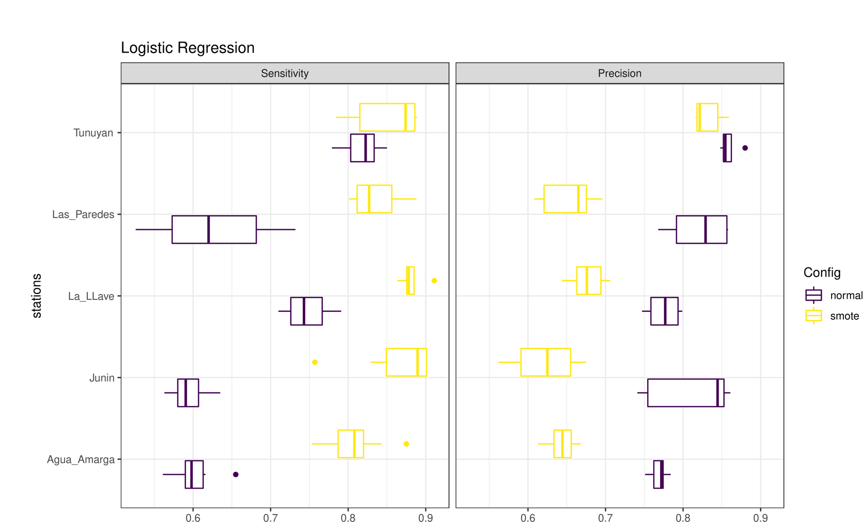

Figure 12:Logistic regression classification: SMOTE vs normal training

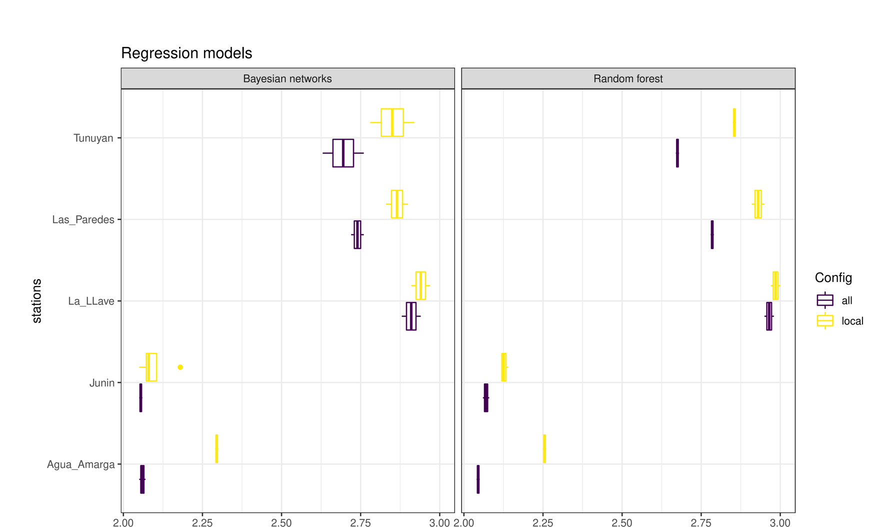

Figure 13:Regression results

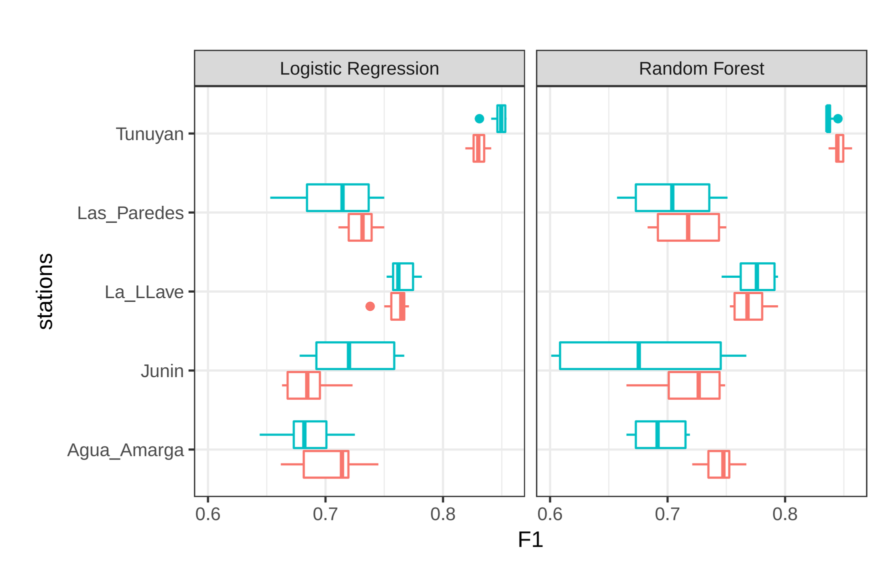

Figure 14:Config all vs local

Conclusions

- In this paper we have created an forecasting engine which is part of an IoT-enabled frost prediction system, which gathers environmental data to predict frost events using machine learning techniques. We have shown that our prediction capability outperforms current proposals in terms of sensitivity, precision and F1.
- The application of SMOTE during the training phase has shown an improved performance in terms of recall in both RF and Logistic Regression models, although the precision decreases.
- We have also observed that, in specific relevant cases, the inclusion of neighbor information helps to improve the precision or recall of the forecasted classification model. On the other hand, regression models have less error by including neighbor information. In these cases, including the spatial relationships, there is a resulting improvement in model performance. We hope to contrast this approach with other scenarios in the future.
- Datasets dacc and dacc-temp perform very well in contrast with dacc-spring, which has data only from Spring season.

Future work

We are working on an R package called frost [1]. The library has implemented the most common empirical methods for frost prediction in agriculture. Shortly, we want to add more sophisticated and accurate models for frost forecasting in agriculture.

References

[1] Ana Laura Diedrichs.
frost: Prediction of minimum temperature for frost forecasting in agriculture, 2018.
R package version 0.0.2. Repository: https://github.com/anadiedrichs/frost.

Institutions

Contact Information

- Web: https://anadiedrichs.github.io/
- Email: ana.diedrichs@frm.utn.edu.ar
- Twitter/Telegram: @anadiedrichs
- www.dharma.frm.utn.edu.ar

Give us feedback!