Interpret Your Care: Predicting the Evolution of Symptoms for Cancer Patients

Rupali Bhati[1], Jennifer Jones[2], Audrey Durand[1,3]

[1]Universite Laval, [2]Princess Margaret Cancer Centre, University Health Network, [3]CIFAR AI Chair

Motivation

- There is a growing population of cancer survivors and high documented rates of acute and persistent treatment effects, e.g., pain, fatigue, but these vary significantly.
- Cancer rehabilitation is becoming increasingly relevant and must be tailored.
- Our goal is to proactively identify the adverse effects of cancer in order to intervene early and provide personalised interventions.

Task

Aim: Predict evolution of a patient's symptoms **(pain or tiredness)** using clinical variables and previously collected patient reported data.

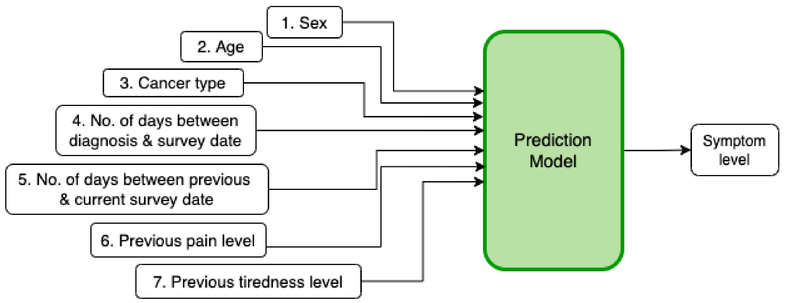
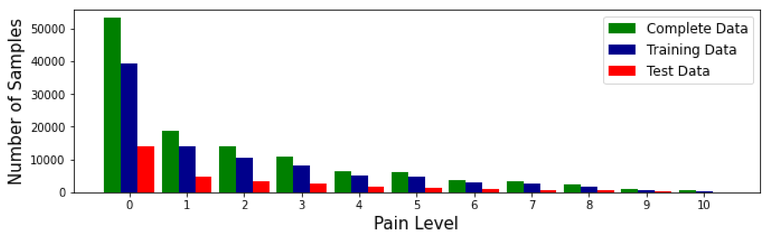
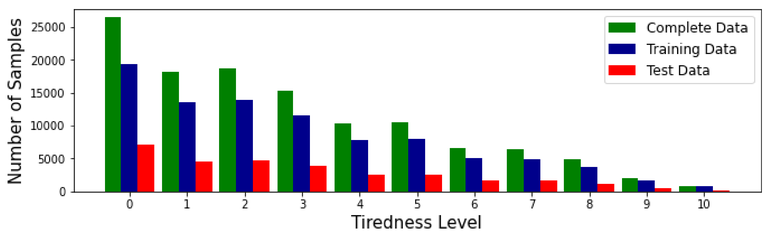


Figure 1: Prediction Model

- Dataset is comprised of real cases collected for the time period of 2013-2019 on **20163** patients with ∼5.95 observations/patient.
- Data is split on date '2017-10-04' and has a train-test split of 75%-25%.
- Symptom level is a discrete value in {0,10}.

Data distribution for patients' pain and tiredness levels*:





*Note the class imbalance.

Methodology

- Comprehensible models → Decision trees based models like LightGBM (Ke et al., 2017).
- **Data imbalance:** We perform oversampling using SMOTE (Chawla et al., 2002).
- **Training:** We do 5-fold cross-validation to determine the maximum tree depth.
- **Performance metrics:** We calculate Mean Absolute Error (MAE) per class as $\text{MAE}_c = \frac{1}{N_c}\sum_{i=1}^{N_c}|y_i - \hat{y}_i|$ and Weighted-MAE to compare different models as $\text{WMAE} = \frac{\sum_{c\in\mathcal{C}} w_c \text{MAE}_c}{\sum_{c\in\mathcal{C}} w_c}$, where $w_c = \frac{\max_{c'\in\mathcal{C}} N_{c'}}{N_c}$ and N_c are no. of samples in class c.

MODELS EVALUATED:

For each symptom, we consider two baselines and two variants of LightGBM models:

- **Naive prior:** Always predicts most frequent symptom level (i.e., 0 for both symptoms).
- **Previous value:** Predicts the current symptom level as the previous symptom level.
- **LP1 and LT1:** This variant relies on clinical variables (i.e., inputs 1 to 5 in Fig.1) and previous level of the predicted symptom (input 6 or 7 in Fig.1) in the prediction. These variants are LP1 and LT1 for pain and tiredness, respectively.
- **LP2 and LT2:** This variant uses all clinical variables in addition to previous levels of both symptoms (all inputs of Fig. 1). We refer to the second variants as LP2 and LT2, for pain and tiredness respectively.

Results

Prediction performance for pain level:

	MAE_0	MAE_1	MAE_2	MAE_3	MAE_4	MAE_5	MAE_6	MAE_7	MAE_8	MAE_9	MAE_{10}	WMAE
NP	0	1	2	3	4	5	6	7	8	9	10	8.72
PV	0.52	0.91	1.3	1.5	1.82	2.04	2.39	2.76	2.89	2.98	5.03	3.88
LP1	0.8	1.34	1.89	2.11	2.3	2.45	2.54	2.67	2.65	2.59	4.41	**3.52**
LP2	0.8	1.35	1.88	2.09	2.26	2.46	2.55	2.72	2.62	2.64	4.53	3.57

Prediction performance for tiredness level:

	MAE_0	MAE_1	MAE_2	MAE_3	MAE_4	MAE_5	MAE_6	MAE_7	MAE_8	MAE_9	MAE_{10}	WMAE
NP	0	1	2	3	4	5	6	7	8	9	10	8.43
PV	0.7	0.99	1.22	1.41	1.55	1.71	1.91	2.06	2.2	2.4	3.25	2.65
LT1	0.84	1.21	1.6	1.86	2.02	2.23	2.29	2.17	2.17	2	2.63	2.33
LT2	0.83	1.21	1.59	1.86	2.01	2.22	2.27	2.18	2.15	2.03	2.52	**2.27**

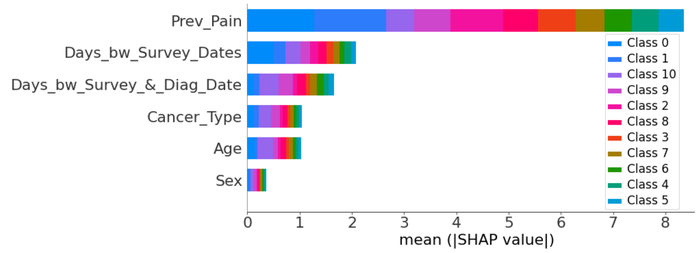
KEY OBSERVATIONS:

- Performance (MAE_c) is inversely proportional to the class density
- Best predictive model for pain level is LP1 (which does not use previous tiredness level)
- Best predictive model for tiredness level is LT2 (which uses previous pain level)
- Not accounting for class imbalance, average of MAE_c's for both NP models is 55/11 = 5.
- WMAE when predicting tiredness is less than WMAE when predicting pain since the former suffers less class imbalance
- LightGBM performs better for higher symptom levels, i.e., MAE_c for c > 6 is higher with PV than with LP1/LP2 and MAE_c for c > 7 is higher with PV than with LT1/LT2.

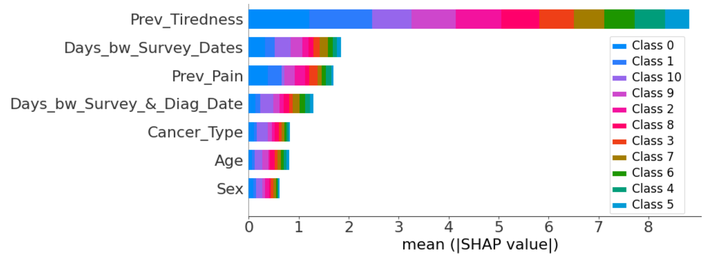
Analysis

We use SHAP values (Lundberg and Lee, 2017) to extract feature importance for the best predictive models of each symptom.

Feature importance for Pain Prediction model LP1



Feature importance for Tiredness Prediction model LT2



KEY OBSERVATIONS:

- There is a clear benefit in leveraging clinical variables in the predictions.
- For both symptoms, the previous symptom levels are the major contributor to the prediction for current symptom levels.
- Previous pain level aids in predicting current tiredness level.

Conclusion

OUR CONTRIBUTIONS:

- Predict patient symptoms using decision trees.
- Use comprehensible evaluation metrics suitable for imbalanced data.
- Showcase that previous symptom levels are important to predict current symptom level.

References

[1] N.V. Chawla et al. "SMOTE: Synthetic Minority Over-Sampling Technique". In: *J. Artificial Intelligence Research, vol. 16, pp. 321-357, 2002.* (2002).

[2] Guolin Ke et al. "LightGBM: A Highly Efficient Gradient Boosting Decision Tree". In: *Advances in Neural Information Processing Systems.* Ed. by I. Guyon et al. Vol. 30. Curran Associates, Inc., 2017.

[3] Scott M Lundberg and Su-In Lee. "A unified approach to interpreting model predictions". In: *Proceedings of the 31st international conference on neural information processing systems.* 2017, pp. 4768–4777.