Model Card for STraTS

Model Details

- Developed by researchers at Virginia Tech, 2022, no information on version update.
- Self-Supervised Transformer for Sparse and Irregularly Sampled Multivariate Clinical Time-Series.
- a Transformer-based architecture called STraTS for clinical time-series, which addresses the unique challenges of missingness and sporadicity of such data by avoiding aggregation and imputation.
- Innovation: A novel CVE mechanism using a one-to-many feed-forward network (FFN) to embed continuous times and measured values to preserve fine-grained information. Utilize forecasting as a self-supervision task to leverage unlabeled data to learn more generalized and robust representations.
- Paper: STraTS

Intended Use

- Intended to be used for healthcare risk prediction of Time-Series data.
- Particularly intended for Healthcare machine learning (HML) practitioners and hospitals.
- Not applicable for anything other than the intended use.

Experimented Factors

- The model showed varied performance across demographic groups.
- Evaluation factor is the highly correlated ethnic attribute. When the model's performance is evaluated against ethnic subgroups (White, Asian, Black, Hispanic/Latino, and others), it shows inconsistent results.

- Further factors like age, gender, insurance, and other covariates are also possible for evaluation.

Metrics

- Evaluation metrics include Receiver Operating Characteristic Area Under the Curve (ROC-AUC) and Precision-Recall Area Under the Curve (PR-AUC).
- Fairness metrics - The 'Fairness Report' provides a detailed report on HML fairness and its metrics.

Training Data

- MIMIC IV v2.0. It is available in the 'Datasheet for CRD'.

Evaluation Data

- 80:20 split for the training and test sets.
- The models were trained for up to 1000 epochs until the validation loss stopped improving for 10 continuous epochs.
- We used the same hyperparameters of the model listed in the code repository.

Model parametrs

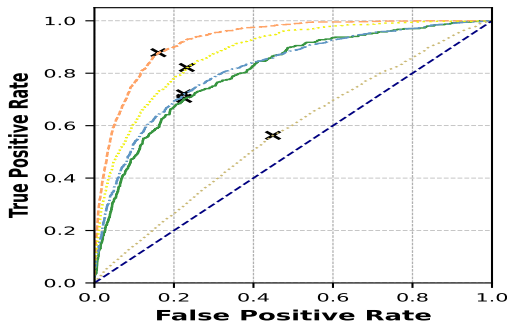- d = 50, M = 2, h = 4, d/o = 0.2, lr = 0.0005. Refer STraTS for further details.

Ethical Considerations

- EHR data is used for evaluating the model's performance where its decision is life-altering based on the intended tasks.
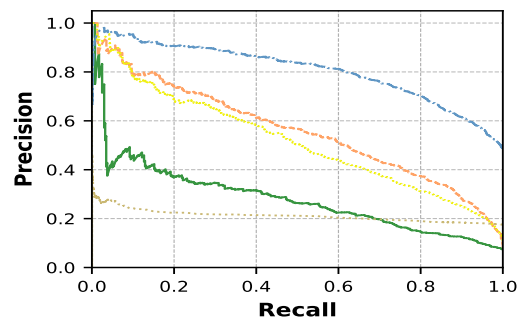
Caveats and Recommendations

- Findings show that the model showed inconsistent results when tested on fairness. It is outperformed by naive baseline XGBoost.

Figure 1: Prediction performance analysis across static and time series models. Panels (a) and (b) show the ROC-AUC and PR-AUC of the models, respectively, and the operating points in (a).