**Applied Machine Learning**

**Final project**

Ortal Parpara 311127005

Yiftah Szoke 204646913

**The article:**

"Deep Incremental Boosting"

**Stage 1- The algorithm:**

1. **A description of the algorithm:**

The algorithm presented in the article is called Deep Incremental Boosting(DIB), which combines the power of AdaBoost, Deep Neural Networks and Transfer of Learning principles, in a Boosting variant which can improve generalization.

1. **The advantages:**

* The algorithm reduces the required training time.
* The algorithm improves generalization.

1. **The disadvantages:**

* The network could be more complexity.

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

**Stage 2- Suggesting an improvement:**

**possible improvements:**

* early stopping
* In future work we will investigate the possibility to modify Deep Incremental Boosting such that only one final test-time Deep Neural Network will be necessary.
* change parameters

**Stage 3- A well-known algorithm for comparison:**

Algorithm for comparison: **Adaboost**(binary) or **Adaboost**.M1(multi-calsses) or **Adaboost.m2**. maybe **bagging**

**Stage 4- Evaluating the algorithms:**

1. **Versions:**

We chose the Boosting Algorithm to evaluate and compare to our improvement.

1. **data:**

Classification task.

**Stages:**

* read 20 datasets (classification).
* data preparation- fill missing data, split the data.

1. **evaluation+** **Hyperparameter Optimization:**

External 10-fold Cross Validation and Internal 3-fold Cross Validation (for optimization).

**Stages:**

* Define the models.
* Create the models using external and internal cross-validation.
* Choose the best parameters for the models.

1. **Performance metrics for evaluation:**

In attached excel.

**Stage 5- Statistical significance testing of the results:**

Choose one of the performance metrics (AUC for example) and use the Friedman test as was presented in the lecture to determine whether the differences are statistically significant. If the results are statistically significant (i.e., the null hypothesis is rejected), do a Post-Hoc test to test the differences between the algorithms. Report and analyze your results and conclusions, <https://www.statology.org/friedman-test-python/>

**Stage 6- Conclusions:**

Present the three algorithms you are evaluating (stages 1-3). Include the pseudo code, results and conclusions.