**Lecture 3: Decision Trees**

**Project 2**

* Να χρησιμοποιήσετε το dataset [**Breast cancer**](https://scikit-learn.org/stable/datasets/index.html#breast-cancer-dataset) και να εφαρμόσετε ένανDecision Tree αλγόριθμο δοκιμάζοντας διάφορες τιμές των παραμέτρων split function (criterion) και max depth. Ο κώδικας να δίνει σαν έξοδο την απόδοση του μοντέλου, όπως αυτή ορίζεται από τις μετρικές Accuracy, Precision, Recall, F1, και να δημιουργεί ένα plot του δέντρου.
* Να εφαρμόσετε τον αλγόριθμο Random Forest δοκιμάζοντας διάφορες τιμές παραμέτρων split function (criterion) και number of trees (n\_estimators). Ο κώδικας να δίνει σαν έξοδο την απόδοση του μοντέλου, όπως αυτή ορίζεται από τις μετρικές Accuracy, Precision, Recall, F1. Επιπλέον, να δημιουργεί 4 γραφήματα (ένα για κάθε μετρική), όπου θα φαίνεται η απόδοση του μοντέλου (άξονας y) καθώς αλλάζει το πλήθος των δέντρων σε ένα μοντέλο Random Forest (άξονας x, από 1 έως 200 δέντρα).

Ενδεικτικά αποτελέσματα από τα πειράματα να καταγραφούν στον παρακάτω πίνακα.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **α/α** | **Algorithm** | **Criterion** | **Max depth** | **Accuracy** | **Precision** | **Recall** | **F1** |
| 1 | Decision Tree |  |  |  |  |  |  |
| 2 | Decision Tree |  |  |  |  |  |  |
| 3 | Decision Tree |  |  |  |  |  |  |
| 4 | Decision Tree |  |  |  |  |  |  |
| 5 | Decision Tree |  |  |  |  |  |  |
| 6 | Decision Tree |  |  |  |  |  |  |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **α/α** | **Algorithm** | **Number of Estimators** | **Criterion** | **Accuracy** | **Precision** | **Recall** | **F1** |
| 1 | Random Forest |  |  |  |  |  |  |
| 2 | Random Forest |  |  |  |  |  |  |
| 3 | Random Forest |  |  |  |  |  |  |
| 4 | Random Forest |  |  |  |  |  |  |
| 5 | Random Forest |  |  |  |  |  |  |
| 6 | Random Forest |  |  |  |  |  |  |