

Github link: <https://github.com/chrisaMel/Machine-Learning.git>

Στο branch : Bankruptcy-Classfier

1. Εισαγωγή

Η παρούσα εργασία επικεντρώνεται στη σύγκριση της απόδοσης διαφορετικών τεχνικών ταξινόμησης για την πρόβλεψη της πτώχευσης επιχειρήσεων. Σκοπός της ανάλυσης είναι να παραθέσει συγκριτικά αποτελέσματα που αποδεικνύουν την ικανότητα των μοντέλων να διαχειριστούν το συγκεκριμένο πρόβλημα, αναδεικνύοντας τις διαφορές στις επιδόσεις τους.

2. Αποτελέσματα και Σύγκριση Μοντέλων

Για την αξιολόγηση της απόδοσης των μοντέλων, χρησιμοποιούνται οι εξής βασικές μετρικές:

- **F1 score:** Υπολογίζει το ποσοστό των σωστά ταξινομημένων περιπτώσεων στο σύνολο των δεδομένων.
- **Πίνακας Σύγχυσης (Confusion Matrix):** Παρουσιάζει τις σωστές και λανθασμένες προβλέψεις για κάθε κατηγορία, παρέχοντας μια πιο λεπτομερή εικόνα της απόδοσης του μοντέλου.

Τα αποτελέσματα δείχνουν ότι:

- Το **Random Forest Classifier** είχε συνολικά την υψηλότερη ακρίβεια, υποδεικνύοντας την ικανότητά του να αναγνωρίζει τα πρότυπα στα δεδομένα.
- Το **SVC (Support Vector Machine)** έδειξε υψηλή ακρίβεια αλλά παρουσίασε δυσκολίες στην επεξεργασία μεγάλου όγκου δεδομένων, με αυξημένο υπολογιστικό κόστος.
- Το **Logistic Regression** είχε τη χαμηλότερη ακρίβεια μεταξύ των αλγορίθμων, υποδηλώνοντας ότι πιθανώς δεν είναι η καλύτερη επιλογή για αυτό το πρόβλημα.
- Το **MLPClassifier (Νευρωνικά Δίκτυα)** παρουσίασε ανταγωνιστικές επιδόσεις, αλλά απαιτεί περισσότερη επεξεργαστική ισχύ και σωστή παραμετροποίηση για να ξεπεράσει άλλα μοντέλα.

3. Συμπεράσματα

Από τη σύγκριση των αποτελεσμάτων, προκύπτει ότι τα μοντέλα δέντρων απόφασης όπως το **Random Forest Classifier** προσφέρουν την καλύτερη ισορροπία μεταξύ απόδοσης και υπολογιστικού κόστους. Αντίθετα, πιο σύνθετες προσεγγίσεις όπως τα νευρωνικά δίκτυα μπορεί να προσφέρουν βελτιωμένες επιδόσεις, αλλά απαιτούν περαιτέρω παραμετροποίηση και αυξημένη υπολογιστική ισχύ.