Τι είναι το διάγραμμα περίληψης SHAP (shap\_summary\_plot)

διάγραμμα περίληψης SHAP, το οποίο είναι ένας συγκεκριμένος τύπος διαγράμματος δύναμης που χρησιμοποιείται για να εξηγήσει την επίδραση διαφορετικών χαρακτηριστικών στην έξοδο ενός μοντέλου μηχανικής μάθησης.

Ακολουθούν οριστικά βασικά χαρακτηριστικά:

* **Σημαντικότητα Χαρακτηριστικών:** Ο άξονας x συνήθως δείχνει τα ονόματα ή τις τιμές των χαρακτηριστικών και η θέση κάθε χαρακτηριστικού στον άξονα υποδεικνύει τη σημαντικότητά του για τις προβλέψεις του μοντέλου. Τα χαρακτηριστικά με μεγαλύτερες θετικές τιμές SHAP συμβάλλουν περισσότερο σε υψηλές εξόδους μοντέλου, ενώ τα χαρακτηριστικά με μεγαλύτερες αρνητικές τιμές SHAP συμβάλλουν περισσότερο σε χαμηλές εξόδους μοντέλου.
* **Κατανομή Επιπτώσεων:** Το χρώμα και το μέγεθος των γραμμών γύρω από τις τιμές των χαρακτηριστικών αντιπροσωπεύουν το μέγεθος και την κατεύθυνση της επίδρασης του χαρακτηριστικού στην έξοδο του μοντέλου. Τα πιο σκούρα χρώματα υποδεικνύουν μεγαλύτερες τιμές SHAP (είτε θετικές είτε αρνητικές). Η εξάπλωση της κατανομής χρώματος γύρω από μια τιμή χαρακτηριστικού υποδεικνύει τη μεταβλητότητα της επίδρασης του χαρακτηριστικού στην έξοδο του μοντέλου. Μια κατανομή με μεγάλη διασπορά υποδηλώνει ότι το χαρακτηριστικό μπορεί να έχει διάφορες επιπτώσεις στην έξοδο του μοντέλου ανάλογα με την συγκεκριμένη περίπτωση.
* **Εξήγηση Μοντέλου:** Τα διαγράμματα περίληψης SHAP είναι ένα πολύτιμο εργαλείο για την ερμηνεία μοντέλων μηχανικής μάθησης, καθώς μπορούν να βοηθήσουν στον εντοπισμό των χαρακτηριστικών που είναι πιο σημαντικά για τις προβλέψεις του μοντέλου και στον τρόπο με τον οποίο αυτά τα χαρακτηριστικά αλληλεπιδρούν για να επηρεάσουν την έξοδο του μοντέλου.