

Εφαρμογή μεθόδων μηχανικής μάθησης για διαχωρισμό σήματος και υποβάθρου σε εξωτικό σενάριο Higgs

Θέλουμε να διερευνήσουμε διάφορες τεχνικές για τη βελτιστοποίηση της επιλογής των γεγονότων σε αναζήτηση υπερσυμμετρικών μποζονίων Higgs στον LHC. Στην υπερσυμμετρία ο Higgs sector αποτελείται από πέντε μποζόνια Higgs σε αντίθεση με το μοναδικό Higgs στο καθιερωμένο μοντέλο. Εδώ εμείς έχουμε ένα βαρύ μποζόνιο Higgs το οποίο διασπάται σε δύο μποζόνια W και ένα τυπικό μποζόνιο Higgs ($H^0 \rightarrow W^+ W^- h$) τα οποία διασπώνται περαιτέρω σε λεπτόνια ($W^\pm \rightarrow l^\pm + \nu$) και b-κουάρκ ($h \rightarrow b \bar{b}$) αντίστοιχα.

Η άσκηση βασίζεται στη δημοσίευση που επισυνάπτεται και η οποία περιέχει πολύ περισσότερες πληροφορίες, όπως γενικές πληροφορίες υποβάθρου, λεπτομέρειες σχετικά με την επιλογή μεταβλητών κλπ. Μπορείτε επίσης να χρησιμοποιήσετε τη δημοσίευση ως οδηγό για τη λύση αυτής της άσκησης.

0: classification (1 = signal, 0 = background)

1 - 21 : low level quantities (var1 - var21)

22 -28 : high level quantities (var22 - var28)

α) Χρησιμοποιήστε τουλάχιστον 2 από τους ταξινομητές classifiers της επιλογής σας για να διαχωρίσετε τα συμβάντα σήματος και παρασκηνίου. Καθορίζοντας την ακρίβεια. (Επιλέγοντας *training set* και *test set* από το ήδη υπάρχων *dataset* θα προσπαθήσετε να κατασκευάσετε μοντέλο πρόβλεψης και να υπολογίσετε την ακρίβειά του).

β) Δοκιμάστε επίσης χρησιμοποιώντας νευρωνικά δίκτυα να υπολογίσετε την ακρίβεια με όποια από τις 2 βιβλιοθήκες (tensorflow-pytorch) επιθυμείτε.

γ) Συγκρίνετε τα αποτελέσματα όταν χρησιμοποιείτε i) τις ποσότητες χαμηλού επιπέδου (low level quantities) και ii) της υψηλού επιπέδου ποσότητες (high level quantities).

Σε αυτά που θα ανεβάσετε θα περιλαμβάνονται, το(α) script(s) της python (.ipynb) και ένα pdf αρχείο που θα επεξηγεί με 2 λόγια τι κάνατε για κάθε ένα μέρος της άσκησης. 1-2 σελίδες το πολύ!