

ВИКОРИСТАННЯ НЕЙРОННИХ МЕРЕЖ ТА NLP ДЛЯ СТВОРЕННЯ СПАМ-ФІЛЬТРУ ДЛЯ ЕЛЕКТРОННОЇ ПОШТИ

Хижниченко Ю.М., Трофімова-Герман А.І.

*Національний технічний університет
«Харківський політехнічний інститут», м. Харків*

Інформаційний простір сучасного світу переповнений різноманітною інформацією, що робить фільтрацію даних нагальною проблемою сьогодення. Зокрема, якісного сортування потребує електронна кореспонденція.

Метою нашої роботи є створення спам-фільтру для україномовних листів на базі NLP та нейронних мереж. Наразі вже існують набори даних для інших мов. Тести іншомовних моделей демонструють точність до 97.7%, що для моделей, котрі не потребують великої кількості даних для навчання є високим результатом. Дослідження включає в себе створення корпусу українських слів та колокацій, притаманних спаму.

Робочим інструментом програми є LDA-класифікатор, що характеризується високою точністю відбору даних. Оскільки, він є алгоритмом навчання з учителем, то необхідні мітки для класів документів. Навчання моделі здійснюється покроково: обчислення центроїду всіх векторів TF-IDF, що входять до класу (наприклад, листи, що є спамом); обчислення центроїду всіх векторів TF-IDF, які не входять до класу (наприклад, листи, що не є спамом), обчислення різниці векторів між центроїдами. Тобто, наближеність нового вектору TF-IDF до певного центроїду визначатиме його приналежність.

Таким чином, за допомогою моделі на базі LDA-класифікатора стане можливим швидко та достовірно визначати, чи є вхідні україномовні повідомлення спамом та сортувати їх відповідним чином.

Література:

1. Linear Discriminant Analysis With Python [Електрон. ресурс]. – Режим доступу: <https://machinelearningmastery.com/linear-discriminant-analysis-with-python/> – Назва з титул. екрану.