

ПОРІВНЯННЯ ШТУЧНОЇ НЕЙРОННОЇ МЕРЕЖІ ІЗ УЗАГАЛЬНЕНОЮ ЛІНІЙНОЮ РЕГРЕСІЙНОЮ МОДЕЛЛЮ ПУАССОНА ПРИ ПРОГНОЗУВАННІ ДАНИХ ПІДРАХУНКУ

Васильєва Н.С., к.ф.-м.н., доцент

(кафедра вищої математики)

Тупко Н.П., к.ф.-м.н., доцент

(кафедра вищої і обчислювальної математики НАУ)

Прогнозування даних підрахунку - одна з ключових задач у страховій галузі, економіці та соціальних науках. Регресійний аналіз зазвичай відноситься до класичного підходу для вирішення цієї задачі. Однак класична регресійна модель Пуассона часто має обмежене застосування, оскільки емпіричні набори даних підрахунку зазвичай демонструють велику дисперсію та надмірну кількість нулів, а отже незбалансованість у даних. Зважаючи на це, а також на позитивні результати машинного навчання у різних галузях, розглянуто його як достойну альтернативу класичному підходу. Проводиться порівняльний аналіз узагальненої лінійної регресійної моделі Пуассона (GLM) з нейронною мережею прямого поширення (Feed Forward Neural Network-FFNN), що є провідним методом машинного навчання, з точки зору прогнозування даних підрахунку і подальшого використання на практиці. Для кращого розуміння моделей, оцінюється їх точність та будуються криві навчання на тестових і навчальних наборах. Крім того, оцінюється важливість вхідних змінних для кращої інтерпретації алгоритмів. Оскільки FFNN є так званим методом «чорної скриньки», для нього не існує прямого способу оцінки змінних. Запропоновано нову технологію оцінки важливості вхідних даних для глибоких нейронних мереж відповідно до принципів теорії інформації. Продемонстровано, що нейронна мережа прямого поширення (FFNN) у порівнянні з узагальненою лінійною регресійною моделлю Пуассона (GLM) забезпечує набагато більшу потужність при незначному збільшенні складності моделі.

Література

1. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press
2. Wüthrich, M.V. (2018). *Data Analytics for Non-Life Insurance Pricing*. ETH Zurich.