

УДК 004

Іванов Д.О.¹, Степаненко О.О.²

¹ студ. гр. КНТ-123м НУ «Запорізька політехніка»

² канд. техн. наук, доц. НУ «Запорізька політехніка»

ДОСЛІДЖЕННЯ ТА ПРОГРАМНА РЕАЛІЗАЦІЯ МЕТОДІВ ОЦІНЮВАННЯ ЗНАЧУЩОСТІ ОПИСОВИХ ХАРАКТЕРИСТИК ОБ'ЄКТІВ

Об'єкт дослідження – процес синтезу розпізнавальних моделей.

Предмет дослідження – методи відбору інформативних ознак.

Мета роботи – дослідження та програмна реалізація методів оцінювання значущості ознак.

Матеріали, методи та технічні засоби: мова програмування Python, середовище розробки PyCharm.

В ході виконання роботи було проаналізовано та досліджено процес синтезу розпізнавальних моделей.

Визначено, що модель є способом опису системи з метою розпізнавання, прогнозування і аналізу стану реальних об'єктів, процесів або явищ; висока складність моделі не завжди гарантує покращення якості розпізнавання, іноді простіші моделі можуть дати кращі результати. Проаналізовано процес моделювання. Відзначено, що важливим етапом в процесі моделювання є визначення змінних і параметрів, на цьому етапі

важливо оцінити інформативність кожної ознак з метою недопущення включення в модель неінформативних ознак або надлишкових ознак. Відбір ознак є процесом зменшення кількості вхідних даних при створенні прогнозної або розпізнавальної моделі, зменшення кількості вхідних змінних може допомогти знизити обчислювальні витрати на моделювання та іноді поліпшити ефективність моделі. Проаналізовано методи відбору ознак. Відзначено, що не існує універсального методу для відбору ознак, тому потрібно знайти найбільш ефективний підхід для конкретної задачі розпізнавання або регресії шляхом систематичних експериментів, необхідно експериментувати з різними моделями, які працюють з різними підмножинами вхідних ознак, відбираючи їх за допомогою різних методів, і визначаючи, який підхід найкраще підходить для конкретної проблеми. Відповідно, для вирішення таких завдань необхідно розробити програмне забезпечення, що дозволяє виконувати оцінювання інформативності ознак за допомогою різних методів.

За результатами проведеного аналізу зроблено висновок, що у наш час існує досить багато програмних засобів для підтримки процесу синтезу розпізнавальних моделей, зокрема, для оцінювання інформативності ознак. Проте деякі програмні засоби для синтезу розпізнавальних моделей можуть не підтримувати широкий спектр методів оцінки інформативності ознак, що може обмежувати можливості аналізу і вибору найбільш релевантних ознак. Іншим недоліком є те, що деякі програмні рішення можуть бути дорогими або вимагати значних обчислювальних ресурсів, що може бути недоступно для невеликих компаній або індивідуальних дослідників. Деяке програмне забезпечення може не завжди ефективно справлятися з незбалансованими наборами даних, де певні ознаки домінують над іншими, що може вплинути на точність оцінки інформативності. Тому актуальною є розробка програмного забезпечення для оцінювання інформативності ознак.

Сформульовано функціональні вимоги до програмного забезпечення для оцінювання інформативності ознак.

Для реалізації програмного забезпечення для підтримки процесу для оцінювання інформативності ознак обрано мову програмування Python, яка має зрозумілий та простий синтаксис, що полегшує написання і читання програмного коду. Python пропонує велику кількість бібліотек для обробки даних і машинного навчання, таких як NumPy, pandas, Scikit-learn, TensorFlow, Keras. Ці бібліотеки надають інструменти для оцінювання важливості ознак, що значно спрощує реалізацію алгоритмів і моделей. Крім того, Python дозволяє швидко тестувати та перевіряти фрагменти коду в інтерактивному режимі, що зручно для експериментів з різними методами оцінювання ознак.

Для створення програмного забезпечення для підтримки процесу оцінювання інформативності ознак обрано середовище розробки PyCharm, що надає потужні інструменти для розробки, налагодження та тестування Python-коду, зокрема, інструменти для автоматичного завершення коду, перевірки синтаксису та інші функції, які полегшують розробку програмного забезпечення. PyCharm підтримує інтеграцію з системами контролю версій, що полегшує співпрацю в командах і керування версіями коду.

Наукова новизна роботи полягає у тому, що створено модель взаємодії користувача з програмним забезпеченням для оцінювання інформативності ознак, яка подана у вигляді діаграми прецедентів, та дозволяє розроблювати програмне забезпечення для відбору інформативних ознак та синтезу розпізнавальних моделей.

Практична цінність результатів роботи полягає у тому, що розроблено програмне забезпечення для оцінювання інформативності ознак.

Запропоновано структуру програмного забезпечення для підтримки процесу оцінювання інформативності ознак. Визначено, що програмне забезпечення для оцінювання інформативності ознак складається з таких основних частин: програмні модулі для оцінювання інформативності ознак (містять програмну реалізацію методів та критеріїв оцінювання інформативності ознак), інтерфейсна частина (містить форми користувача для роботи з програмою), програмні модулі для роботи з вибірками даних (містять засоби для роботи з файлами, що містять вибірки вхідних даних).

Описано особливості реалізації програмного забезпечення для оцінювання інформативності ознак. Виконано проєктування інтерфейсу взаємодії користувача з програмним забезпеченням для підтримки процесу оцінювання інформативності ознак.

Виконано тестування розробленого програмного забезпечення для оцінювання інформативності ознак. Результати тестування програмного забезпечення показали, що розроблена програма дозволяє забезпечити автоматизацію процесів, пов'язаних з підтримкою процесу оцінювання інформативності ознак.

Галузь використання – розпізнавання образів, моделювання.