

Програмні засоби для аудіодетекції звукових сигналів немовленнєвого походження

Вінницький національний технічний університет

Анотація. Розглядаються програмні засоби класифікації звукових сигналів із використанням нейронних мереж. Проводиться огляд наявних алгоритмів класифікації звуку.

Ключові слова : штучні нейронні мережі, аудіосигнали, алгоритми, технології машинного навчання, метод опорних векторів, логістична регресія, метод k-найближчих сусідів, модель Гаусової суміші, наївний Баєсів класифікатор, дерева рішень.

Abstract. Software tools for classification of sound signals using neural networks are considered. An overview of available sound classification algorithms is conducted.

Keywords: artificial neural networks, audio signals, algorithms, machine learning technology, support vector machine, logit model, K-nearest neighbors algorithm, Gaussian Mixture Models, Naive Bayes Classifier, decision tree.

Сьогоднішня кількість наявної аудіоінформації стрімко зростає. Як результат, зростає необхідність в розвитку систем, спрямованих на обробку цієї інформації в автоматизованому режимі, без залучення людей. Однією з важливих завдань є класифікація звукових сигналів. Наразі існує досить велика кількість алгоритмів, що спрямовані на вирішення цих задач, проте дослідження теми лішається актуальним.

Технології машинного навчання в цілому — це один з найцікавіших і найбільш ефективних способів вирішення подібних задач. З часів створення перших обчислювальних машин, люди мріяли створити машину, яка буде навчатися та розв'язувати задачі. Ця мрія призвела до розвитку цілої галузі науки, відомої зараз як наука про штучний інтелект.

Штучні нейронні мережі наразі — один із найбільш популярних і вживаних розділів штучного інтелекту, що показав свою високу результативність на багатьох задачах, таких як розпізнавання зображень, машинний переклад, розпізнавання мови, та інших, не менш складних. Якісне розв'язання цих задач ще кілька десятиріч років тому вважалося вкрай важким, до моменту стрімкого вибухового розвитку глибоких штучних мереж. Дослідження їх можливостей та підходів, що базуються на таких мережах — також є вкрай важливим з точки зору подальшого розвитку науки, адже це сприяє їх подальшому розвитку.[4]

Класифікація - один з розділів машинного навчання, присвячений розв'язанню наступної задачі: в наявності множина об'єктів (ситуацій), розділених деяким чином на класи[1]. Задана певна обмежена множина об'єктів, для яких відомо, до якого класу вони відносяться. Ця множина називається навчальною вибіркою. Класова приналежність інших об'єктів не відома. Вимагається побудувати алгоритм, здатний класифікувати будь-який об'єкт початкової множини

Класифікувати об'єкт — значить віднести об'єкт до певного класу, встановити між ними відповідність.

Для класифікації звуку може бути використана більшість алгоритмів машинного навчання, які здатні розв'язувати задачу класифікації.

Метод опорних векторів (англ. support vector machine, SVM), є машинним алгоритмом, котрий навчається на прикладах та використовується для класифікації об'єктів[2]

Наприклад, SVM може розрізнити аварійний режим роботи електромеханічної системи та класифікувати його за наявності попередніх досліджень, можливих за технологічними вимогами режимів роботи. Такий підхід розкриває значні можливості для побудовання адаптивних систем автоматичного керування. В основі SVM лежить деяка математична сутність – алгоритм максимізації деякої математичної функції відносно наявного набору даних.

Логістична регресія або логіт-регресія (англ. logit model) – це статистична модель, що використовується для передбачення ймовірності виникнення деякої події шляхом підгонки даних до логістичної кривої. Також вона може бути визначена як регресійна модель, що передбачає

категоріальну змінну[1]. Незважаючи на назву, цей вид регресії використовується переважно для класифікаційних, а не регресійних задач.

Метод k -найближчих сусідів[1] (англ. K-nearest neighbors algorithm, k -NN) - метричний алгоритм для автоматичної класифікації об'єктів або регресії. У разі використання методу для класифікації об'єкту присвоюється той клас, який є найбільш поширеним серед k сусідів даного елемента, класи яких вже відомі. Можлива модифікація, яка полягає в присвоєнні більшої ваги до класу найближчих сусідів та меншої — до більш віддалених.

Модель Гаусової суміші (англ. Gaussian Mixture Models) є імовірнісною моделлю, яка базується на припущенні, що всі спостереження, що належать до певних класів, утворюються з суміші скінченної кількості гаусових розподілів з невідомими параметрами [1]. Модель сумішей можна розглядати як узагальнення кластеризації k -means з включенням інформації про коваріаційну структуру даних і центри розподілів.

В цілому GMM модель цікава тим, що розглядає дані як результат лінійної комбінації певної кількості гаусових розподілів, в той час, як більшість інших методів статистичної класифікації використовують чи лише один певний розподіл, чи взагалі розглядають лише взаємовідносини між даними.

Наївний баєсів класифікатор — ймовірнісний класифікатор, що використовує теорему Баєса для визначення ймовірності приналежності спостереження (елемента вибірки) до одного з класів при припущенні (наївному) незалежності змінних. [3]

Деревя рішень належать до самих популярних і потужних інструментів, що дозволяють вирішувати задачі класифікації. В основі роботи дерев рішень лежить процес рекурсивної розбивки вхідної множини спостережень або об'єктів на підмножини, асоційовані із класами.

Висновок: Технології машинного навчання на сьогоднішній день є основним способом розпізнавання звукових сигналів. Були розглянуті алгоритми машинного навчання, які можуть використовуватись для розв'язання задач розпізнавання звуків.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Trevor H. The elements of statistical learning. Second edition / H. Trevor, R. Tibshirani., 2008.
2. Шеремет О. Метод опорних векторів / О. Шеремет, В. Садовой. // Мат. Мод. № 1. – 2013. – №28. – 13 с.
3. Domingos P. On the optimality of the simple Bayesian classifier under zeroone loss / P. Domingos, P. Pazzani. // Machine Learning. – 1997. – №29. – 103–137 с.
4. Николенко С. И. Глубокое обучение. Погружение в мир нейронных сетей. / С. И. Николенко, А. А. Кадурын, Е. О. Архангельская., 2018. – 480 с

Тютюнник Ярослав Александрович, ст. гр. 2KI-15b факультет інформаційних технологій та комп'ютерної інженерії. Вінницький національний технічний університет, м. Вінниця, 2ki15b.tiutiunnik@gmail.com.

Чорний Денис Сергійович, ст. гр. 1KI-15b факультет інформаційних технологій та комп'ютерної інженерії. Вінницький національний технічний університет, м. Вінниця, 1ki15b.chornyi@gmail.com.

Науковий керівник: Ткаченко Олександр Миколайович канд. техн. наук, доцент кафедри обчислювальної техніки, Вінницький національний технічний університет, м. Вінниця, alextk1960@gmail.com.

Yaroslav Tiutiunnyk, student, 2KI-15b, Faculty of information Technologies and Computer Engineering, Vinnytsa National Technical University, Vinnytsia, 2ki15b.tiutiunnik@gmail.com.

Denys Chornyi, student, 1KI-15b, Faculty of information Technologies and Computer Engineering, Vinnytsa National Technical University, Vinnytsia, 1ki15b.chornyi@gmail.com.

Supervisor: Olexsandr Tkachenko PhD, assistant professor of the department of Computer Technology, Vinnytsia National Technical University, Vinnytsia, alextk1960@gmail.com.