

УДК 004.522

АВТОМАТИЧНА ІДЕНТИФІКАЦІЯ ЛЮДИНИ ЗА ГОЛОСОМ З ВИКОРИСТАННЯМ ШТУЧНОЇ НЕЙРОННОЇ МЕРЕЖІ

А.В. КИНКУРОГОВА^{1*}

¹*студентка факультету електроніки, НТУУ «КПІ», Київ, УКРАЇНА*

^{*}*kynkurogova.a@gmail.com*

Голос – така ж невід’ємна риса кожної людини, як і його обличчя або відбитки пальців. Широке поширення засобів зв'язку (стаціонарні та мобільні телефонні мережі, інтернет-телефонія і т.д.) відкривають великі можливості для застосування даного ідентифікатора. Крім того, розпізнавання по голосу досить зручно для користувачів і вимагає від них мінімум зусиль.

Задача ідентифікації диктора за голосом полягає у автоматичному визначенні особи по мовному сигналу. Можна виділити наступні типи даної задачі: «Закрита» вибірка – всі диктори апріорі відомі системі розпізнавання, необхідно визначити якому саме з відомих дикторів належить той чи інший голос; «Відкрита» вибірка – необхідно визначити чи належить голос одному з апріорі відомих дикторів або невідомому диктору.

Загалом, вирішення задачі ідентифікації людини по голосу складається з трьох основних етапів: визначення класифікаційних ознак, створення і навчання моделей або класифікатора, тестування.

В якості класифікаційних ознак, як правило використовуються частота основного тону, фонемі, мел-кепстральні коефіцієнти, коефіцієнти лінійного передбачення. З використанням вищевказаних ознак для кожного диктора будується своя модель на основі закону їх розподілу або створюється загальний класифікатор [1].

База даних та класифікаційні ознаки.

База даних складалася з 83 дикторів, для кожного з яких було записано 5 фрагментів мовних сигналів з частотою дискретизації 8 кГц та глибиною квантування 16 біт. Таким чином, загальний об'єм даних складав 415 записів довжиною від 5 до 27 с кожен.

В якості класифікаційних ознак були обрані мел-частотні кепстральні коефіцієнти, які знаходяться наступним чином [1]:

- сигнал розбивається на фрейми довжиною 25 мс;
- для кожного фрейму знаходиться спектральна щільність потужності за допомогою дискретного перетворення Фур'є;
- спектральна щільність потужності розбивається на смуги за допомогою мел-частотних вікон і обчислюється сумарна енергія сигналу в кожній смузі;
- отримані енергії логарифмуються;
- застосовується дискретне косинусне перетворення Фур'є.

На практиці використовують від 2 до 13 отриманих таким чином коефіцієнтів.

Штучні нейронні мережі.

Нейронні мережі – штучні математичні моделі, а також їх програмна та апаратна реалізація, побудовані за принципом функціонування біологічних нейронних мереж (мереж нервових клітин живого організму). Система, архітектура і принцип дії базується на аналогії з мозком живих істот. Ключовим елементом цих систем виступає штучний нейрон як імітаційна модель нервової клітини мозку – біологічного нейрона [2].

На рис. 1 представлена найпросторіша тришарова нейронна мережа, що складається з вхідного шару, прихованого шару та вихідного шару.

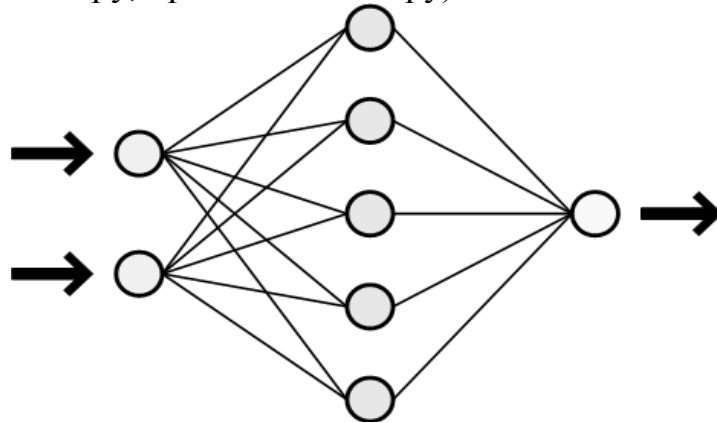


Рис. 1 – Схема простої нейронної мережі

У сфері штучного інтелекту, штучні нейронні мережі були успішно застосовані для розпізнавання мови, аналізу зображень та адаптивного управління, для того, щоб побудувати так званих програмних агентів (в комп'ютерних та відео іграх) або автономних роботів [3].

На етапі навчання вищеописаної нейронної мережі був використаний алгоритм перехресної перевірки, суть якого криється в тому, що загальний об'єм даних розбивається на k частин, після чого $k-1$ частини використовуються для навчання класифікатора, а остання частина – для тестування. Вказана процедура повторюється k разів.

На етапі тестування нейронної мережі була отримана точність 95%, яка співрозмірна з результатами інших дослідників. В роботі показано, що штучні нейронні мережі можуть бути використані для вирішення задачі ідентифікації диктора.

Список літератури:

1. Coleman J. *Introducing speech and language processing*. Cambridge, UK: Cambridge University Press. – 2005.
2. Fausett L. *Fundamentals of neural networks: architectures, algorithms, and applications*. – 1994.
3. Вікіпедія – вільна енциклопедія [Електронний ресурс]. – <http://wikipedia.org>.