

АВТОМАТИЗОВАНА СИСТЕМА ЗБОРУ ДАНИХ ДЛЯ МАШИННОГО НАВЧАННЯ

Автор Ялосоветський В. І.

(науковий керівник — к.т.н., доц. Мосійчук В. С.)

Виявлення захворювань за допомогою нейронних мереж є перспективним напрямом медичної діагностики, особливо у разі скринінг досліджень [1]. Проте ефективність методів машинного навчання є ефективною лише у разі наявності значної вибірки накопичених даних як здорових пацієнтів, так і з патологіями. Крім того статистична обробка медичних досліджень базується на принципі того, що те що вірне для випадкової вибірки – вірно і для генеральної сукупності (популяції), з якої ця вибірка отримана. Проте вибрати або набрати істинно випадкову вибірку з генеральної сукупності на практиці є задачею складною, та вимагає проведення спеціалізованих досліджень. Тому слід прагнути, щоб вибірка була репрезентативною по відношенню до досліджуваної популяції для того, щоб вона найбільш адекватно описувала всі можливі особливості досліджуваного стану або захворювання в певній групі [2]. Оскільки отримати з відкритих джерел такі дані важко, а їх релевантність для конкретної задачі може бути відносною, то важливим є створення системи та інфраструктури для швидкого набору бази даних цільової групи пацієнтів з повним описом їх функціонального стану. У цьому контексті важливим для машинного навчання є автоматизація як процесів реєстрації біосигналів, так і їх збереження з повним описом метаданих.

Шляхи зменшення ручної роботи оператора як предмет автоматизації

У разі використання звичайного медичного обладнання, зазвичай, автоматизація накопичення даних є відсутньою. Дані накопичуються безпосередньо в діагностичній апаратурі. Далі відповідно необхідно виконувати ряд маніпуляцій для структурування та збереження даних на комп'ютері з подальшою конвертацією біосигналів та вилучених з них параметрів, збереження результатів до бази даних.

Вирішенням проблеми може стати створення системи та інфраструктури для швидкого та автоматизованого набору бази даних, зокрема у вигляді безпосереднього збереження необроблених даних з сенсорів на сервері.

Для цього необхідно стає розробка ряду нових пристроїв реєстрації біосигналів з можливістю прямого доступу до сервера і баз даних. Такий підхід відповідає загальній тенденції, що пов'язана з розвитком технологій Інтернету речей та хмарних сервісів.

Приклад реалізація системи

Для розробки автоматизованої системи збору даних доцільно реалізувати прилад-сенсор на основі Wi-Fi модуля ESP-8266 [3] та сенсора для детектування пульсової хвилі MAX30100 [4].

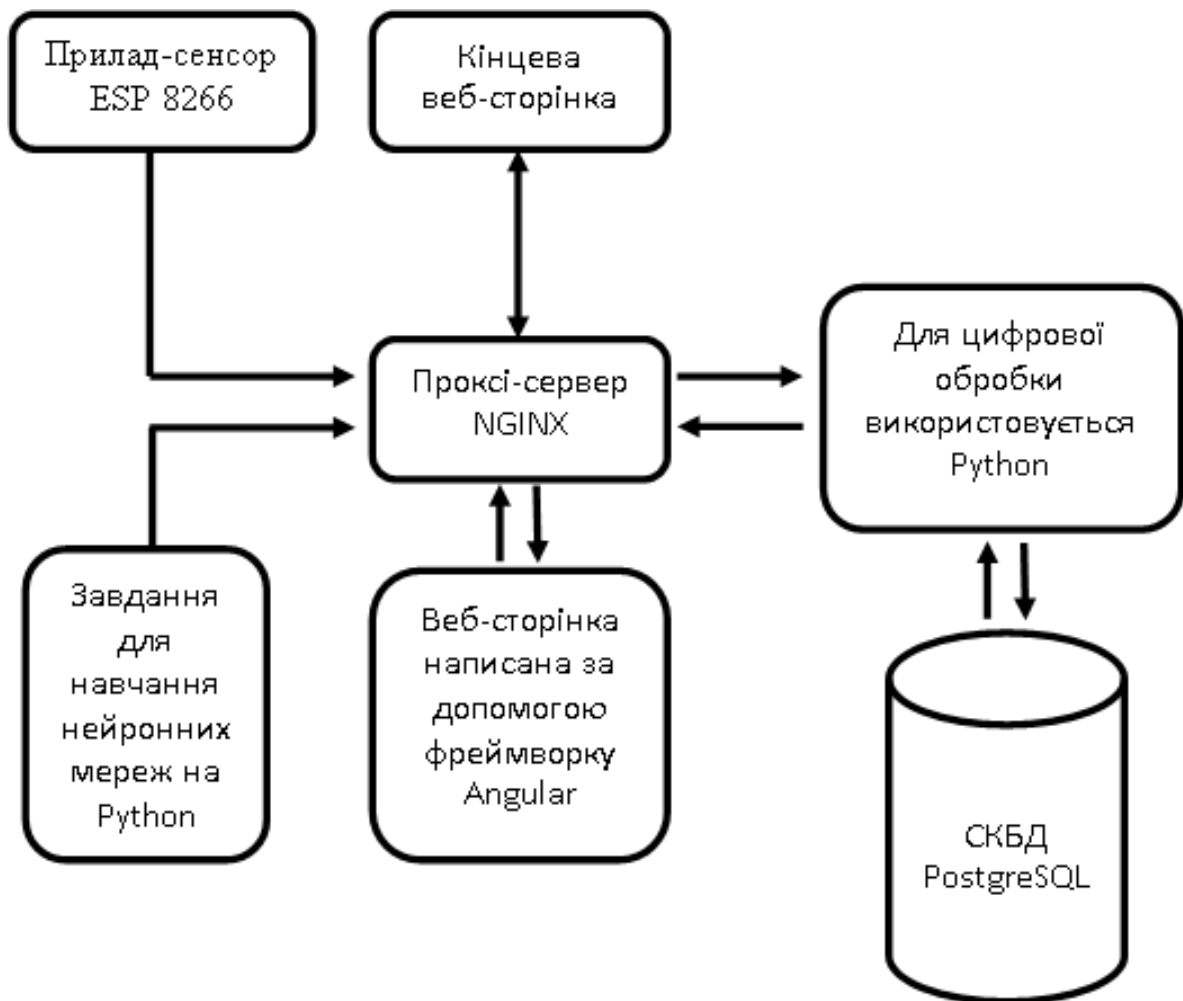


Рисунок 1. Структурна схема побудованої системи

База даних зроблена на PostgreSQL. Сам веб-ресурс розроблений на базі Angular та бібліотеки PrimeNG. В якості мови програмування для навчання нейронних мереж, фільтрації та обробки пакетів використано Python3. Мова Python дозволяє швидко створювати прототипи, не втрачаючи високу ефективність за рахунок попередньої компіляції в байт-код і значної частини стандартної бібліотеки, написаної на C [5]. Тому Python в більшості випадків виконує код швидше переважної більшості динамічних мов програмування (PHP, Ruby, JavaScript). Крім того на Python вже існують рішення у вигляді бібліотек для навчання нейронних мереж та роботи з масивами TensorFlow, PyTorch, NumPy, Keras, Theano, Matplotlib, SciPy [6]. Для детектування хвороб є сенс користувати спеціалізовану для медичних даних нейронну мережу ANFIS(Adaptive Neuro-Fuzzy Inference System). [7]. Це адаптивна нечітка нейронна мережа, а висновок робиться на основі апарату нечіткої логіки, ці функції налаштовуються за допомогою алгоритму зворотного поширення помилки.

Автоматичної реєстрації пульсових хвиль та вилучення параметрів

В процесі розробки приладу-сенсора виявилось, що для можливості передачі великих масивів даних до сервера, дані треба відправляти кількома пакетами з подальшим їх об'єднанням на сервері.

На рис. 2 можна подивитися на реалізацію інтерфейсу з можливістю автоматичного набору бази даних цільової групи пацієнтів з описом їх функціонального стану. На сервері автоматично ведеться логування та присвоєння унікального ідентифікатору записів.

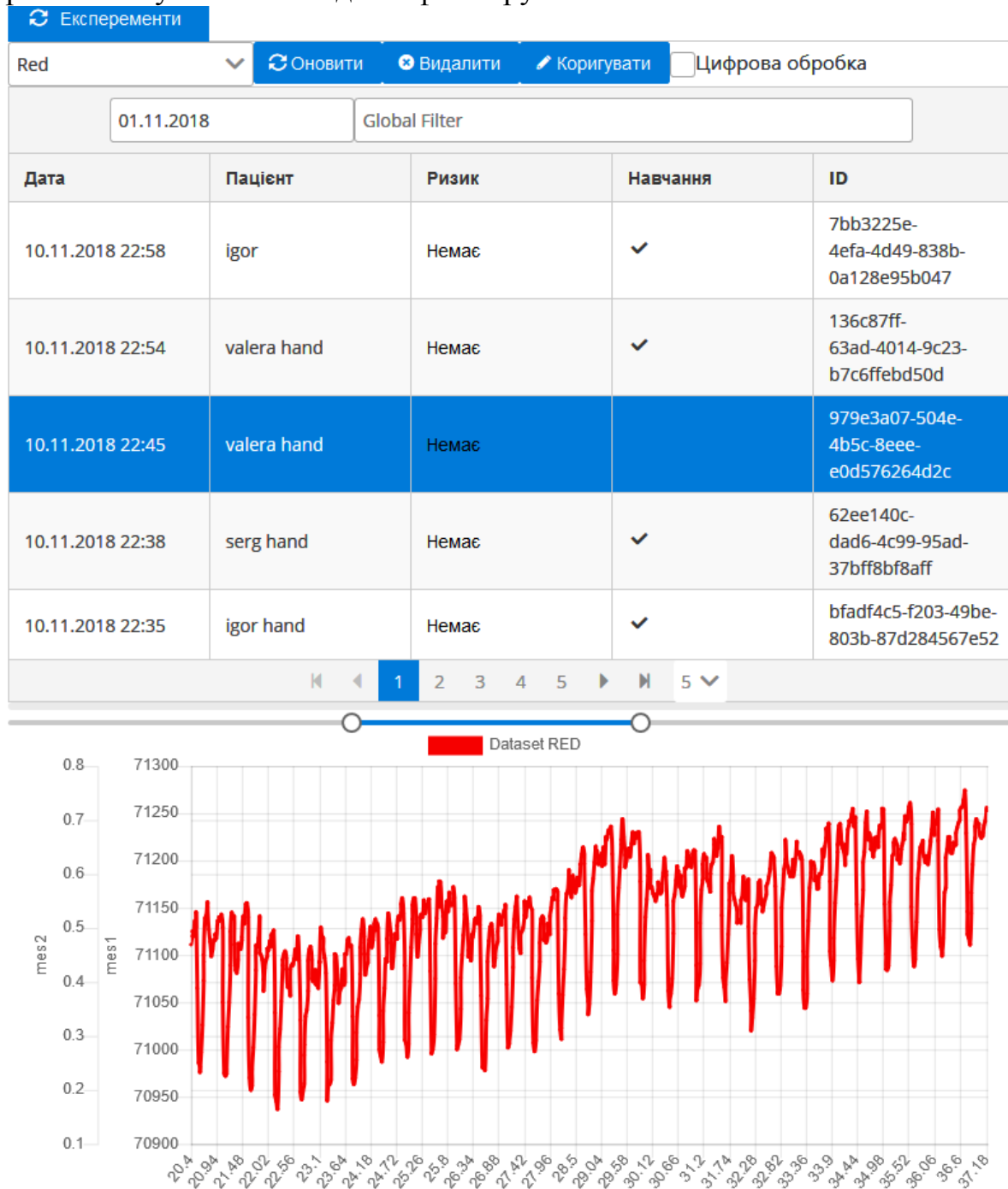


Рисунок 2. Приклад сигналу з порушеною структурою

Висновки

Для побудови напівавтоматичного пристрою-сенсора збору діагнозів хворих та їх функціонального стану доцільно розробляти систему яка містить серверну частину для збереження даних з різних пристроїв до однієї бази даних. Окрім цього на сервері повинні бути підготовлені служби для навчання нейронних мереж. На даний час найзручніше працювати з нейронними мережами на мові програмування Python. В системі повинна бути інтерфейс для коригування та вводу інформації для опису функціонального стану пацієнтів.

Перелік посилань

1. Агаджанов М. Нейросети диагностируют проблемы с сердцем более точно, чем врачи <https://habr.com/post/370621/>
2. Кочетов А. Г. Методы статистической обработки медицинских данных: Методические рекомендации для ординаторов и аспирантов медицинских учебных заведений, научных работников / А. Г. Кочетов, О. В. Лянг, В. П. Масенко, И. В. Жиров, С. Н. Наконечников, С. Н. Терещенко, – М. : РКНПК, 2012. – 42 с.
3. ESP8266EX Datasheet [Електронний ресурс]. – Режим доступу: https://www.espressif.com/sites/default/files/documentation/0a-esp8266ex_datasheet_en.pdf
4. MAX30100 Pulse Oximeter and Heart-Rate Sensor IC for Wearable Health [Електронний ресурс]. – Режим доступу: <https://datasheets.maximintegrated.com/en/ds/MAX30100.pdf>
5. Каштамов А. Почему Python? [Електронний ресурс]. – Режим доступу: <https://khashtamov.com/ru/why-python/>
6. Fabisch A. Библиотека нейронных сетей с открытым исходным кодом для глубокого обучения [Електронний ресурс]. – Режим доступу: <http://qaru.site/questions/104486/open-source-neural-network-library>
7. Безруков Н. С. Построение и моделирование адаптивной нейро-нечеткой системы в задаче медицинской диагностики / Н. С. Безруков, Е. Л. Еремин // Информатика и системы управления. – 2005. – Вип. 2. – с. 36-46.
8. Применение цифровой обработки сигналов / под ред. Э.Оппенгейма. – М. : Мир, 1980. – 552 с.

Анотація

Основний зміст дослідження становить аналіз та загальні вимоги до системи автоматизованого збору даних для подальшого навчання по цим даним нейронних мереж. Показано основні принципи побудови системи автоматичної реєстрації та збору метаданих про функціональний стан.

Ключові слова: нейронні мережі, база даних, автоматизація збору даних, реєстрації біосигналів.

Abstract

The main content of this research is the analysis and general requirements for the system of automated data collection for further study on these data of neural networks. In this work, the author showed the basic principles of constructing a system of semi-automatic collection of functional state of patients and elements of their further analysis.

Keywords: neural networks, Database, automatic data collection, detector biosignal.