

В. О. Новіков¹
С. М. Злепко²
К. С. Навроцька²

ПЕРСПЕКТИВА ЗАСТОСУВАННЯ AZURE MACHINE LEARNING В МЕДИЧНИХ ДОСЛІДЖЕННЯХ

¹Херсонський національний технічний університет;

²Вінницький національний технічний університет

Розглянуто приклад застосування математичної моделі машинного навчання в постановці діагнозу гемограми. За результатами оцінки моделі зроблено висновок.

Ключові слова: гемограма, Machine Learning, математична модель.

Вступ

Сьогодні все більшою популярністю користуються сервіси постановки онлайн-діагнозу за симптомами, проте в дійсності ці сервіси ніяк не можуть замінити лікаря, повністю так як безумовно фахівець враховує багато факторів і має певний особистий досвід.

Разом з тим, з'являються програмні сервіси, які можуть якщо не замінити людину, то принаймні дублювати його професійну діяльність [1].

Так, існує сервіс компанії Microsoft хмарних обчислень, який здатний на основі побудованої математичної моделі навчатися і прогнозувати результат певної події. Наприклад, цей сервіс використовується для відстеження шахрайських операцій з банківськими картами, за своєю суттю модель навчається на типових операціях і потім фіксує аномальні транзакції. Цей сервіс цілком можливо використовувати для потреб охорони здоров'я, тому що сервіс Microsoft Cortana Intelligence Suite передбачає застосування і в цій галузі.

У роботі розглянута можливість застосування такого сервісу Azure Machine Learning в лабораторній діагностиці. Як приклад, для побудови математичної моделі постановки попереднього діагнозу обрано метод гемографічного дослідження крові.

У разі застосування звичайного програмування обробка результату обмежувалася б рамками допустимих мінімумів і максимумів параметрів гемограми, що вводяться. Те що входило б в межі, — норма, за межами, — патологія. Відповідно результат носив би рекомендаційний характер, тому що необхідно враховувати безліч другорядних факторів таких як додаткові медичні процедури пацієнта, які можуть впливати на результат дослідження. Наприклад, озонотерапія може привести до панцитопенії — різкого зниження вмісту еритроцитів, тромбоцитів, лейкоцитів крові [2]. Всі ці фактори враховувати в програмному коді досить-таки складно.

Застосування ж машинного навчання дозволяє додати штучному інтелекту відсутній досвід. Таким чином можна аналізувати всі дані, якими оперує лікар, а, як приклад, використанням кореляційного аналізу можна встановити взаємозв'язок змінних.

Метою роботи є створення моделі, яка зможе класифікувати наявність або відсутність патології в гемограмі.

Методика експерименту

Для цього використана двійкова класифікація, за допомогою якої можна отримати відповідь (так чи ні), тобто присутня патологія чи ні. Для постановки конкретного діагнозу (визначити тип патології) треба було б використовувати мультикласову класифікацію. Для цього для навчання моделі необхідна більша кількість інформації. Під час побудови математичної моделі використані такі модулі: ініціалізація моделі, класифікація, навчання, прогноз, оцінювання.

Основна частина

За результатами обстеження пацієнтів протягом певного періоду (для порівняння взято дані за один день і один тиждень) створена база даних. База даних, виконана в форматі CSV, містить такі

дані: ПБ, стать, вік, показники гемограми: кількість еритроцитів, гемоглобіну, колірного показника, ретикулоцитів, гематокрит, швидкість осідання еритроцитів, тромбоцити, лейкоцити, еозинофіли, базофіли, нейтрофіли, лімфоцити, моноцити, дані про наявності / відсутності патології, діагноз.

Після розміщення бази даних, за допомогою модуля `select column's in dataset` вибирають ті дані, які можуть безпосередньо впливати на визначення статусу наявності/відсутності патології. Вибираємо всі параметри крім ПБ, діагнозу. Після розбивки даних за допомогою модуля `split` проводиться навчання отриманої моделі. Потім додається блок оцінки моделі та результату. На цьому етапі оцінюється точність моделі (можна довіряти її даним чи ні), приклад діалогового вікна зображено на рис. 1.

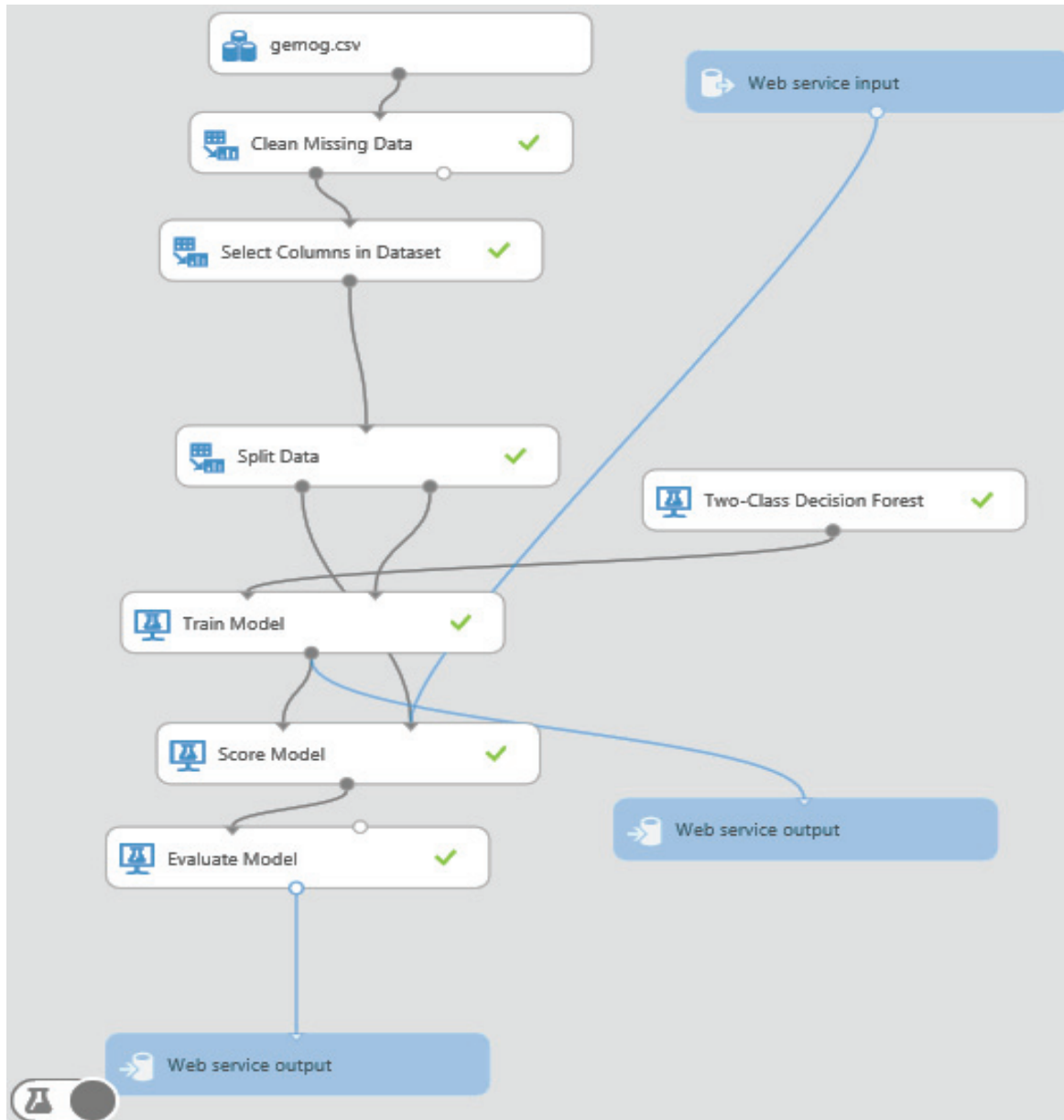


Рис. 1. Зразок діалогового вікна

Рис. 2 демонструє можливість порівняння точності моделей з даними пацієнтів за один день і один тиждень. У першому випадку, як видно, результату моделі довіряти не можна, у другому випадку модель навчилася на більшому масиві даних і її дані ближче до значення 1, тобто довіряти її результатам можна.

Як альтернативу в побудові моделі, для виявлення грубих даних і встановлення патологій за ними, можна було б використовувати модуль виявлення аномалій.

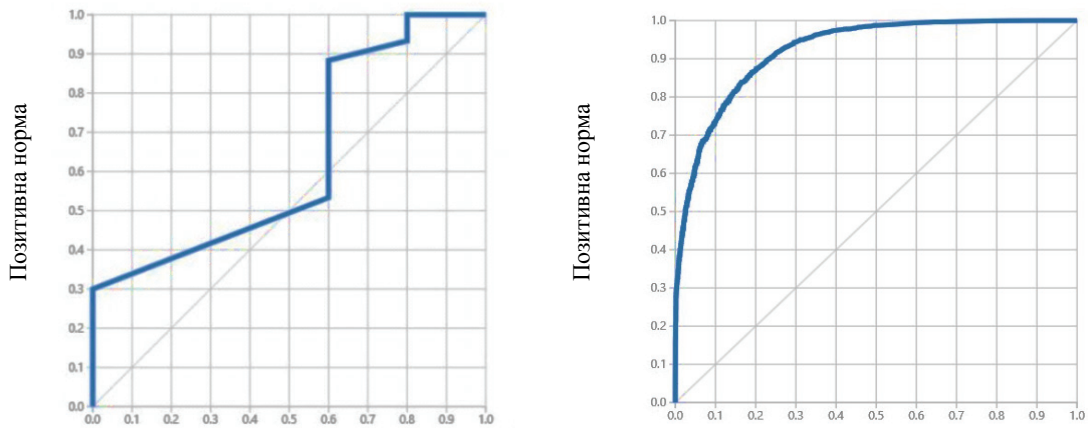


Рис. 2. Порівняння точності моделей

За результатами навчання моделі створюється Web API для загального використання сервісу (рис. 3).

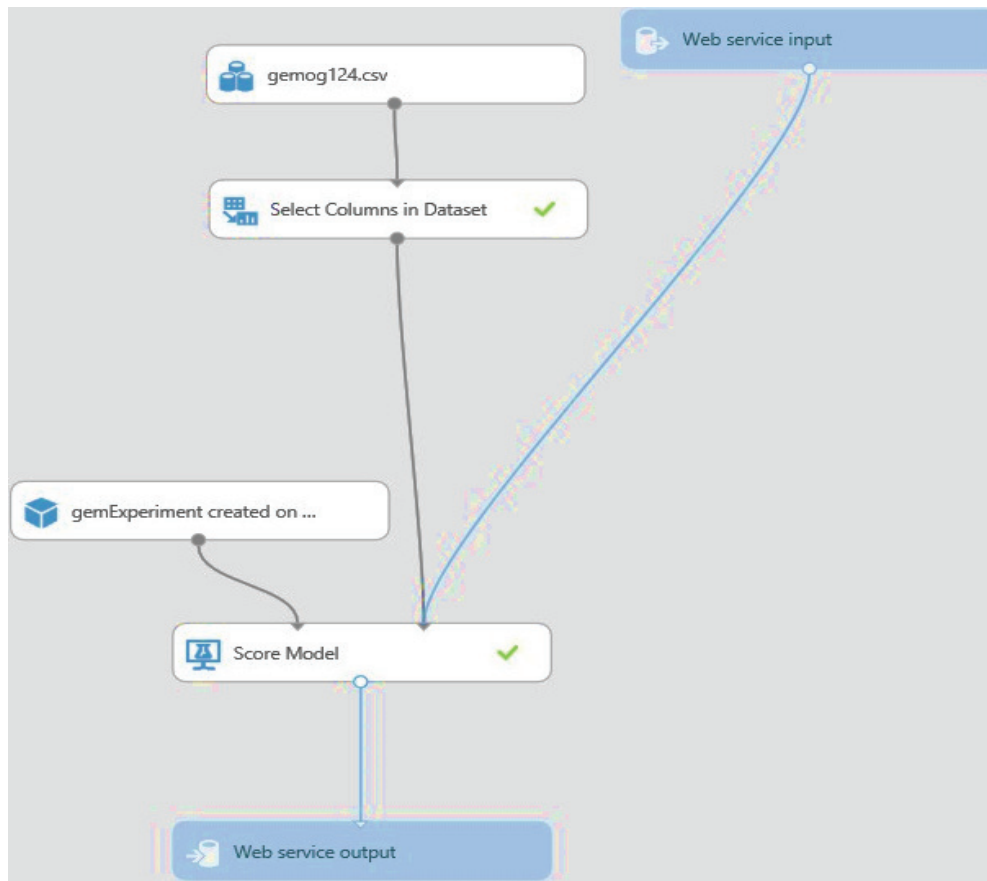


Рис. 3. Діалогове вікно з модулями Web API

Проведено апробацію моделі шляхом введення даних патології. Результат видається моделлю «Паталогія»:

```
Result:{"Results":{"output1":{"type":"table","value":{"Column Names":["Patalog","Er","Gemog","Color","Reticuloc","Gematocrit","Trombocit","SOE","Leic","Aosin","Bazo","Limfoz","Momoz","Scored","String","Int32","Int32","Int32","Int32","Int32","Int32","Int32","Int32","Int32","Int32","Int32","Int32","String","Double"],"Values":[[null,"4","125","1","1","0","200","10","5","3","1","2","5","Yes","0,625"]]]}}
```

Висновки

Результати дослідження дозволяють стверджувати, що використання сервісу хмарних обчислень машинного навчання для постановки проміжного діагнозу є цілком можливим. Перевагою цього сервісу, є те що він може працювати з великими масивами інформації за рахунок використання хмарних обчислень, і залежить від обладнання клінічних лабораторій.

У перспективі штучний інтелект зможе давати точніший прогноз, ніж людина, оскільки він може базуватися на великому досвіді (наприклад, якщо розробити і використовувати всесвітню базу даних пацієнтів). Виключається також людський фактор.

Крім того, сервіс може в перспективі спростити роботу з графічною обробкою інформації, наприклад, під час роботи з фаціями шляхом оцифрування зображення і визначення за математичною моделлю по карті глибин наявності/відсутності патології [3].

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Markatou Xu H. Machine learning and word sense disambiguation in the biomedical domain: design and evaluation issues / Xu H. Markatou, M. R. Dimova, Liu. H. Friedman // BMC bioinformatics. — 2006. — No. 7 (1). — P. 14—20.
2. Bocci V. A. Archives of Medical Research / V. A. Bocci // Scientific and Medical Aspects of Ozone Therapy. — 2006. — No. 37 (4), — P. 425—435.
3. Novikov V. Potential influence of wireless Wi-Fi networks for the digestive function of a stomach / V. Novikov, A. Borsuk, V. Glazkova // 13th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET). IEEE. — 2016.

Рекомендована кафедрою проектування медико-біологічної апаратури ВНТУ

Стаття надійшла до редакції 4.10.2016

Новіков Всеволод Олександрович — канд. техн. наук, старший викладач кафедри інформаційно-вимірювальних технологій електроніки й інженерії, e-mail: vsevolodnovikov@live.com.

Херсонський національний технічний університет, Херсон;

Злепко Сергій Макарович — д-р техн. наук, професор, завідувач кафедри проектування медико-біологічної апаратури e-mail: smzlepko@ukr.net;

Навроцька Ксенія Сергіївна — аспірантка кафедри проектування медико-біологічної апаратури, e-mail: ksysha33@ukr.net.

Вінницький національний технічний університет, Вінниця

V. O. Novikov¹
S. M. Zlepko²
K. S. Navrotska²

Perspective of Azure Machine Learning in Medical Research

¹Herson National Technical University;

²Vinnitsia National Technical University

There has been considered the example of a mathematical model of machine learning in diagnosis of hemogram. The evaluation model has been concluded.

Keywords: hemogram, Machine Learning, mathematical model.

Novikov Vsevolod O. — Cand. Sc. (Eng.), Senior Lecturer of the Chair of Information and Measuring Technology and Electronics Engineering, e-mail: vsevolodnovikov@live.com;

Zlepko Sergii M. — Dr. Sc. (Eng.), Professor, Head of the Chair of Design of Biomedical Equipment e-mail: smzlepko@ukr.net;

Navrotska Ksenia S. — Post-Graduate Student of the Chair of Design of Biomedical Equipment, e-mail: ksysha33@ukr.net

В. А. Новиков¹
С. М. Злепко²
К. С. Навроцкая²

Перспектива применения Azure Machine Learning в медицинских исследованиях

¹Херсонский национальный технический университет;

²Винницкий национальный технический университет

Рассмотрен пример применения математической модели машинного обучения в постановке диагноза гемограммы. По результатам оценки модели сделан вывод.

Ключевые слова: гемограмма, Machine Learning, математическая модель.

Новиков Всеволод Александрович — канд. техн. наук, старший преподаватель кафедры информационно-измерительных технологий электроники и инженерии, e-mail: vsevolodnovikov@live.com;

Злепко Сергей Макарович — д-р техн. наук, профессор, заведующий кафедрой проектирования медико-биологической аппаратуры e-mail: smzlepko@ukr.net;

Навроцкая Ксения Сергеевна — аспирантка кафедры проектирования медико-биологической аппаратуры, e-mail: ksysha33@ukr.net