



Енергетика

Секція: Енергетика

УДК 004.4

SEMI-SUPERVISED MACHINE LEARNING FOR OIL AND GAS PIPELINE CRASHES DETECTION

НАПІВКЕРОВАНЕ МАШИННЕ НАВЧАННЯ ДЛЯ ВИЯВЛЕННЯ НЕСПРАВНОСТЕЙ НАФТОГАЗОПРОВОДІВ

Kropyvnytska V.B. / Кропивницька В.Б.

s.t.s., as.prof. / к.т.н., доц.

ORCID: 0000-0001-5231-7104

Mahas D.M. / Магас Д.М.

*post-graduate student / аспірант**Ivano-Frankivsk National Technical University of Oil and Gas,**Ivano-Frankivsk, Karpatska, 15, 76019**Івано-Франківський національний технічний університет нафти і газу,**Івано-Франківськ, вул. Карпатська, 15, 76019*

Анотація. В роботі розглядається використання машинного навчання для виявлення несправностей нафтогазопроводів. Виділяються основні види машинного навчання. Більш детально описується процес напівкерованого (semi-supervised) машинного навчання в контексті нафтогазової промисловості. Зазначаються його переваги та недоліки.

Ключові слова: нафтогазопроводи, машинне навчання, відмови, трубопровід.

Вступ.

Транспортування є одним з головних напрямків нафтогазової промисловості поряд з розвідкою, бурінням та видобуванням. Одним з найбільш ефективних способів доставки нафти та газу є трубопровідний транспорт. Він відрізняється відносно меншою загрозою аварійних ситуацій (наприклад у порівнянні з залізничним транспортом). При цьому, у разі виникнення аварійної ситуації, шкода довіллію може бути надзвичайно великою, зокрема через складність виявлення аварійних ділянок та усунення наслідків. Сучасні напрацювання у сфері цифрових технологій дозволяють спростити моделювання та фактичну роботу нафтогазових трубопровідних систем. При цьому часто використовуються реальні операційні дані з роботи трубопроводу, які вимагають чіткої кластеризації та математичного моделювання, для продукування найбільш коректного результату. Ці дані можуть бути використаними для визначення проблемних ділянок трубопроводу.

Основний текст.

Застосування штучного інтелекту дозволяє обійти часткові обмеження у створенні моделей представлення для складних фізичних систем, зробити їх більш передбачуваними та контрольованими. Надзвичайно важливим при цьому є вид машинного навчання, що використовується:

- з учителем (supervised);
- без вчителя (unsupervised);
- напівкероване (semi-supervised).



При навчанні з вчителем, спостереження за системою та дані, які є результатом цього спостереження, використовуються для встановлення залежності між певними характеристиками системи та наперед визначеними виходами. Сам набір даних сформований таким чином щоб «навчити» алгоритми класифікувати дані та передбачати результат з максимальною точністю. Така модель може бути використаною для передбачення результату роботи для нових елементів системи (напр. трубопроводів) з таким же набором характеристик. Навчання з вчителем вирішує проблеми двох типів:

- класифікації;
- регресії.

Класифікація передбачає використання алгоритмів з метою розподілення даних між окремими категоріями. Прикладом може бути встановлення типу аварійної ситуації. Завданням регресії є встановити зв'язки між залежними та незалежними змінними. Регресійні моделі дають змогу передбачати числові значення на основі вхідних даних. Прикладом може бути визначення впливу робочого тиску трубопроводу на термін його експлуатації.

Навчання без вчителя не опирається на наявність наперед визначених результатів та специфічного набору тренувальних даних. Машинне навчання без вчителя передбачає вирішення наступних задач:

- кластеризації;
- асоціації;
- зменшення розмірності.

Кластеризація передбачає віднесення окремих процесів та сутностей на основі спільної ознаки до певного кластеру. Схожість процесів та сутностей визначається на основі стратегії некерованого (unsupervised) навчання та способу тренування. Результуюча модель використовується для співставлення нових даних з наперед визначеними кластерами. Асоціація передбачає використання різноманітних правил для знаходження зв'язку між змінними у наборі даних, що розглядається. Зменшення розмірності використовується коли кількість вхідних даних є надто великою. Зміст полягає у зменшенні кількості вхідних даних до допустимого розміру, зберігаючи при цьому їх цілісність.

Напівкероване навчання поєднує окремі елементи навчання з та без вчителя з метою покращення швидкодії моделі та/або її точності. При напівкерованому навчанні окрім звичних даних виду «вхід-вихід» також застосовуються неповні дані або дані без чіткого означення «виходу». Для підтримки неповних даних використовуються кластеризаційні методи, що навчаються на неповних даних та роблять ці дані придатними для навчання з вчителем.

Різнорізані підходи керованого машинного навчання використовуються для класифікації та регресії з метою адресування конкретних цілей моделювання трубопроводних систем. Прикладами таких підходів є: дерева класифікації та регресії (CART), штучні нейронні мережі (ANN), нечітка логіка (FL), модель поверхні реагування (RSM), метод опорних векторів (SVM) [1], тощо. Для прикладу, деякі праці описують застосування ANN для передбачення стану морських нафтогазопроводів, на основі історичних даних [2]. Серед



характеристик трубопроводу, які використовувались для передбачення стану трубопроводу, були вік, робочий тиск, зношення металу та діаметр труби. Особливої уваги заслуговує праця [3], що описує розробку кількох регресійних та ANN моделей для передбачення причини відмов у нафтопроводах. Ключовим диференціатором виступають типи відмови: механічні, спричинені корозією або операційні. Хоча моделі такого типу є надзвичайно ефективними у визначенні причини відмови, дані на яких базується навчання не є загальнодоступні та самі по собі є доволі дорогими для продукування.

В нафтогазовій індустрії існує брак кооперації в сфері машинного навчання, що традиційно базується на наявності великої кількості даних у вільному доступі та обміні професійним досвідом. Потенційним рішенням проблеми може бути використання публічно доступних баз даних, що містять звіти про інциденти для широкого спектру трубопроводів. В більшості випадків такі бази даних формуються на основі попередніх звітів про інцидент та деталізуються згідно з їх розвитком. Внесення у журнал відмов є обов'язковим у випадку будь якої несправності, тому збір цих даних є простішим та більш ефективним з фінансової сторони як для державних органів регулювання так і для постачальників послуг[4].

Використання баз даних на основі звітів про відмови є доволі перспективним з позиції прискорення вирішення потенційних проблем, але тільки за умови подолання обмежень притаманних для таких баз даних, зокрема відсутності інформації. Брак інформації перешкоджає формуванню вхідних даних, які підійшли б для навчання з вчителем. Ймовірним рішенням може бути використання напівкерованих методів машинного навчання, що підтримують не тільки кластеризацію та класифікацію, а й безперервний етап форматування вхідних даних також відомий як «імпутація даних».

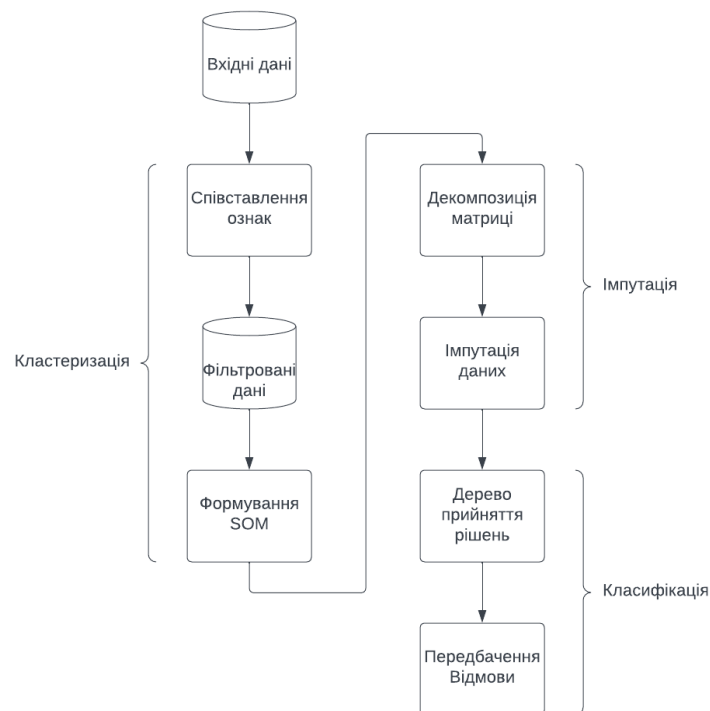


Рисунок 1 – Процес передбачення відмови на основі напівкерованого машинного навчання



Опис процесу «кластеризація-імпутація-класифікація» (СІС) вже знайшов своє відображення у наукових публікаціях [5]. Запропонований фреймворк доповнює дослідження оптимальної кластеризації за допомогою методики факторизації. При цьому кластеризація використовується не тільки для прискорення процесу класифікації як у СТС («кластеризуй потім класифікуй») фреймворках, а й спрямовує формування вибірки корисної інформації для ефективної факторизації з метою імпутації відсутніх даних. Узагальнений процес може бути представлений так як це зображено на рис. 1.

Висновки.

Було розглянуто основні види машинного навчання в контексті нафтогазової промисловості. Зазначено важливість використання відкритих баз даних з інформацією про інциденти на підприємствах. До переваг напівкерованого машинного навчання можна віднести: можливість працювати з неоднорідними та частково відсутніми даними (зокрема з безкоштовних відкритих джерел), гнучкість механізму навчання при правильній реалізації, доволі високу точність передбачень. Серед недоліків слід зазначити значно складнішу реалізацію ніж у випадку навчання з або без вчителя.

Література:

1. Ani, M., Oluyemi, G., Petrovski, A., Rezaei-Gomari, S. // SPE Intelligent Energy International Conference and Exhibition. (OnePetro, 2016).
2. El-Abbasy, M. S., Senouci, A., Zayed, T., Mirahadi, F. & Parvizsedghy, L. // Artificial neural network models for predicting condition of offshore oil and gas pipelines. Autom. Constr. 45, 50–65 (2014).
3. Senouci, A., Elabbasy, M., Elwakil, E., Abdrabou, B. & Zayed, T. // A model for predicting failure of oil pipelines. Struct. Infrastruct. Eng. 10, 375–387 (2014).
4. Kabir, G., Sadiq, R. & Tesfamariam, S. // A fuzzy Bayesian belief network for safety assessment of oil and gas pipelines. Struct. Infrastruct. Eng. 12, 874–889 (2016).
5. H. Alobaidi, M., Meguid M., Zayed, T. // Semi-supervised learning framework for oil and gas pipeline failure detection

***Abstract.** The article covers the usage of machine learning for oil and gas pipelines failure detection. It describes main types of machine learning with a main focus being on semi-supervised ML in oil and gas industry.*

***Key words:** oil and gas pipelines, machine learning, failures, pipelines.*

Науковий керівник: к.т.н., доц. Кропивницька В.Б.

Стаття відправлена: 13.03.2023 г.

© Магас Д.М.