МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

<<ЛЬВІВСЬКА ПОЛІТЕХНІКА>>

Інститут ІКНІ

Кафедра систем штучного інтелекту



ЗВІТ

Лабораторна робота **№5**

З курсу “ Інтелектуальний аналіз даних”

Виконав:

**Бурак Марко**

гр. **КН-310**

Прийняв(ла):

Якимишин Х.М.

Львів – 2020

**Лабораторна №5**

**Класифікація часових рядів методами Orange**

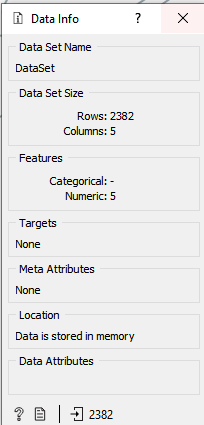
На основі супутникових даних перевірити гіпотезу, що поява піків сплесків в характеристиках сонячного вітру (bulk speed) може спричинити появу піків на графіках швидкості вітру ураганів.

В якості навчальних даних (training set) вибрати дані по урагану Irma. А на основі цих даних побудувати прогноз ураганів Jose та Katia (Dev Set)

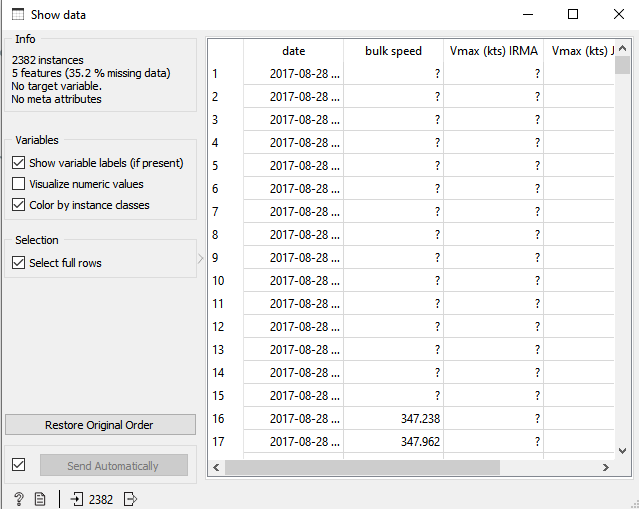
Дані для аналізу знаходяться у файлі DataSet.xslx

За допомогою Orange:

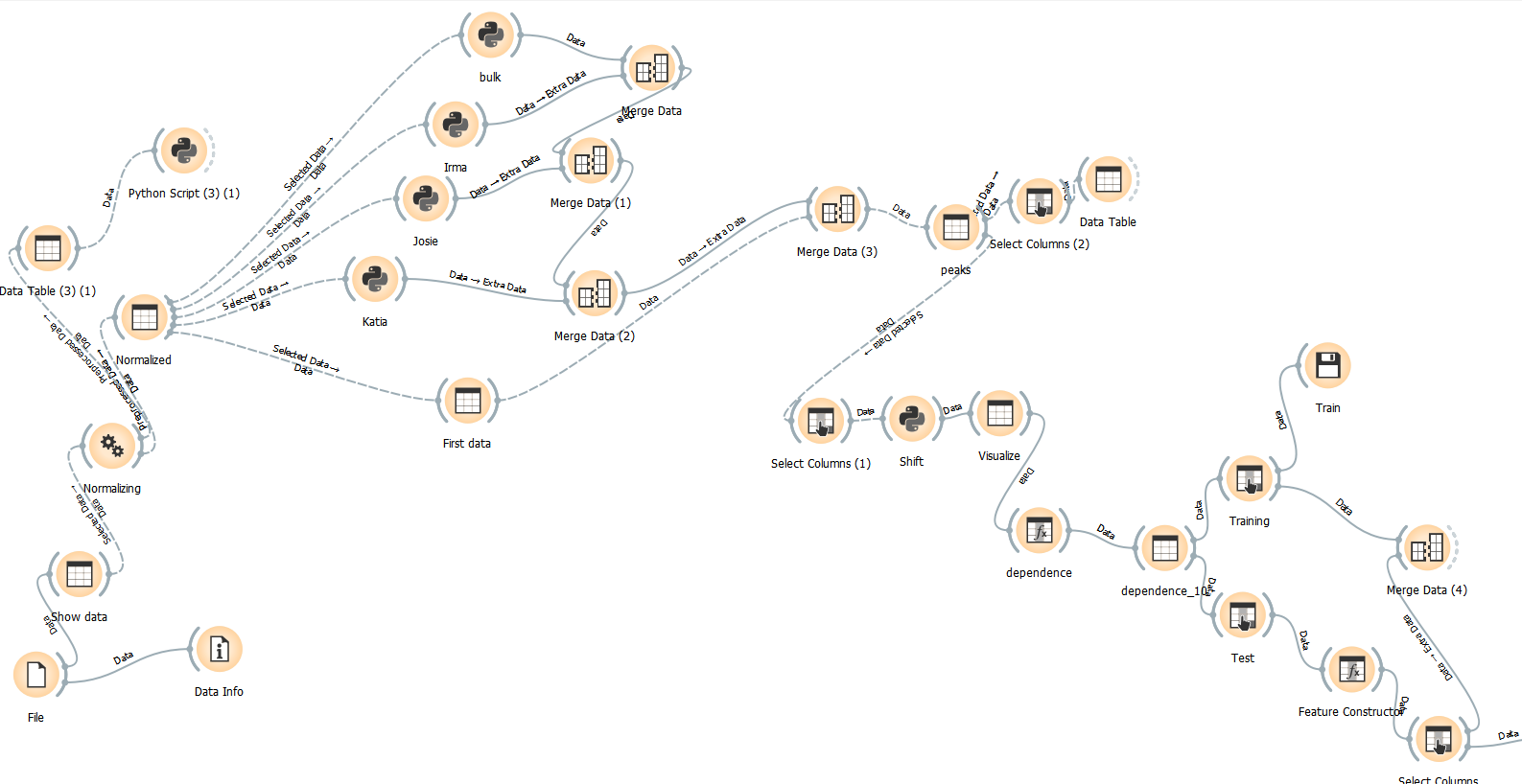
1. Нормалізувати дані навчального набору
2. Написати віджет для пошуку піків в часових рядах даних (на вході часовий ряд, на виході бінарний часовий ряд 0 – відсутній пік, 1 – є пік)
3. Перетворити часові ряди за допомогою зробленого віджету в бінарний вигляд
4. Лагово трансформувати часовий ряд bulk speed для врахування лагів із затримкою 10 год.
5. Перевірити гіпотезу, що піки на часовому ряді bulk speed за попередні 10 год можуть спричинити появу піків на рядах даних швидкості вітру ураганів. Тобто на вхід класифікатора подаються піки за 10 годин на bulk speed. На виході – наявність піку на швидкості вітру урагану.
6. Для перевірки класифікатора вибрати Test&Score. Training set – Дані по урагану Irma. Test Set – об’єднані дані для ураганів Jose та Katia. В якості класифікаторів перевірити всі наявні в Orange стандартні класифікатори.
7. Побудувати  Confusion Matrix для аналізу точності прогнозів та вибрати 4 кращі класифікатори.
8. Оформити аналіз отриманих даних у вигляді короткого звіту.



Інформація про підвантаженій датасет

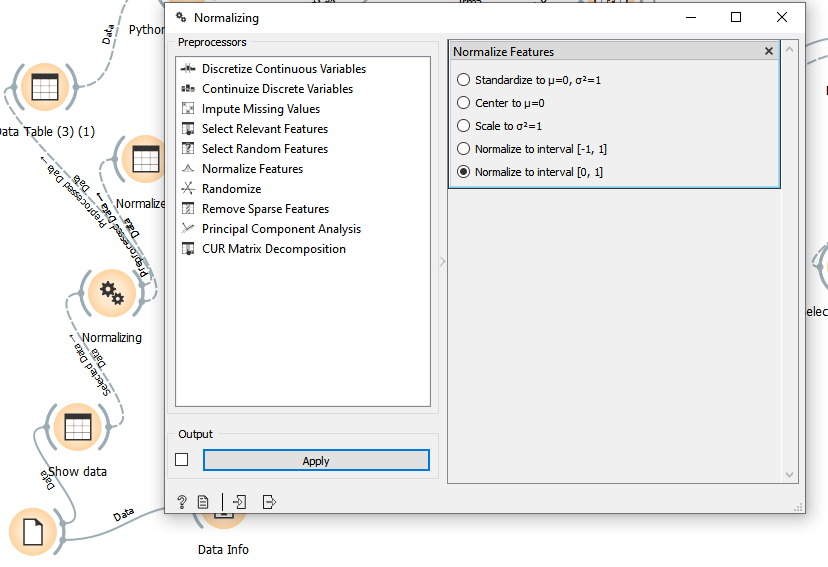


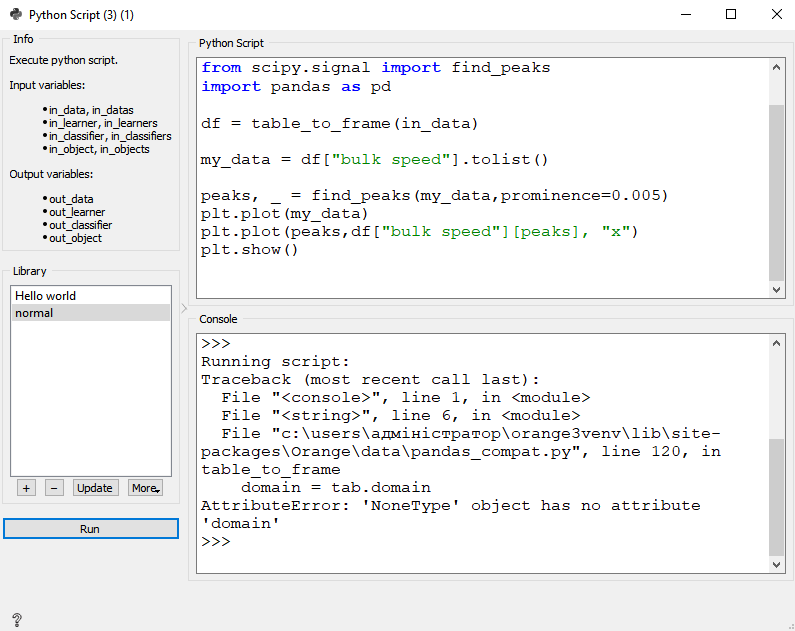
Дані початкові

****

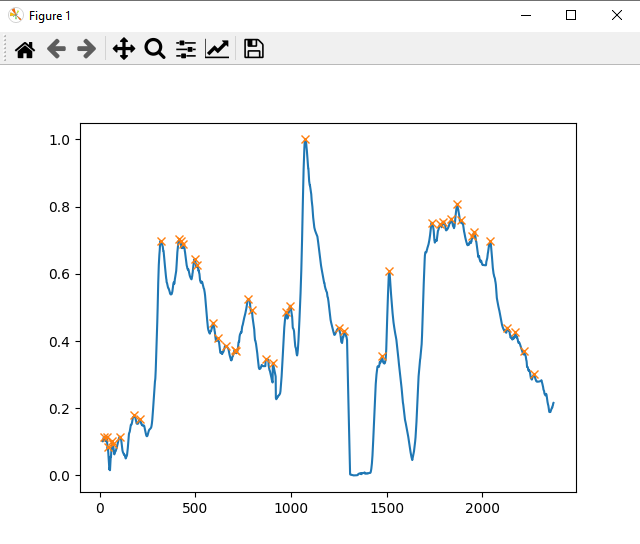
Проект

**Нормалізація даних**



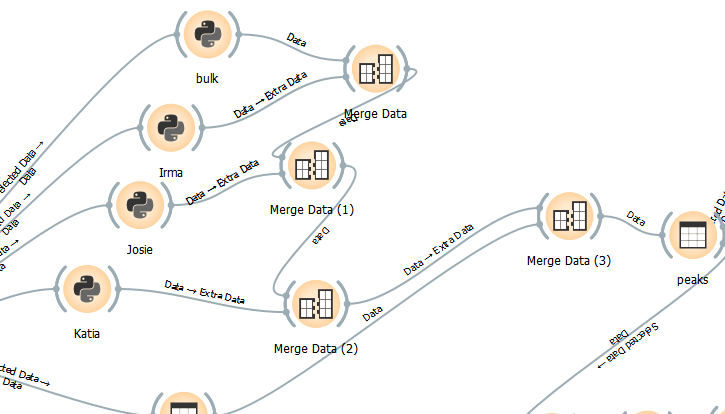
****

Срипт для графіку



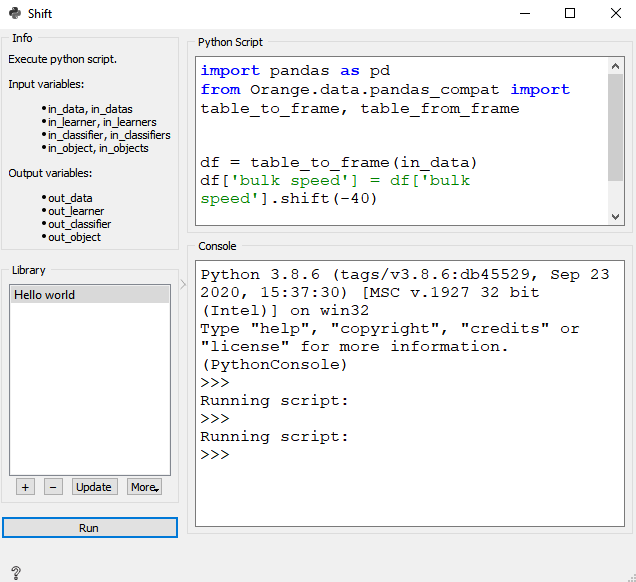
Піки

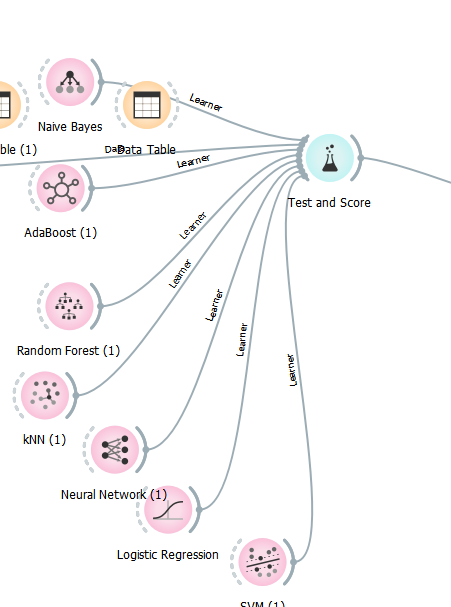
Написати віджет для пошуку піків в часових рядах даних (на вході часовий ряд, на виході бінарний часовий ряд 0 – відсутній пік, 1 – є пік)



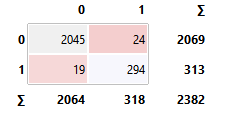
Віджет для піків

Лагово трансформувати часовий ряд bulk speed для врахування лагів із затримкою 10 год.





Класифікатори та тест скор



Конфюжн матрікс

**Висновок:** виконуючи лабораторну роботу, провів аналіз даних та за допомогою класифікаторів побудував конфюжн матрікс у якій 2045 знайдено правдивих 0 і 294 правдивих 1, а неправдивих лише 19 і 24, отже програма працює правильно.