

Інформаційні системи планування авіаперельотів компанії Ryanair

Виконала: Чечельницька Олександра Борисівна

Науковий керівник: д.т.н., проф. Зайченко Олена Юріївна

Постановка задачі

- Провести порівняльний аналіз методів прогнозування ARIMA та Штучних Нейронних Мереж
- Створити програмний продукт, який буде давати прогноз цін на авіаквитки
- Проаналізувати отримані результати, порівняти їх зі справжніми даними

Об'єкт дослідження

- Об'єктом дослідження обрано дані про авіарейси компанії Ryanair за 2016-2018 рік, а саме – дата, координати аеропорту вильоту, координати аеропорту призначення, тривалість польоту, клас, валюта та опублікована ціна квитка

Предмет дослідження

- Предметом дослідження є алгоритм розрахунку ціни на авіаквитки з використанням даних авіакомпанії, прогнозованих за допомогою нейронних мереж та авторегресії з ковзним середнім

Мета роботи

- Побудувати ефективну оцінку вартості квитка на літак, що, в свою чергу, дозволить максимізувати дохід бюджетної авіакомпанії

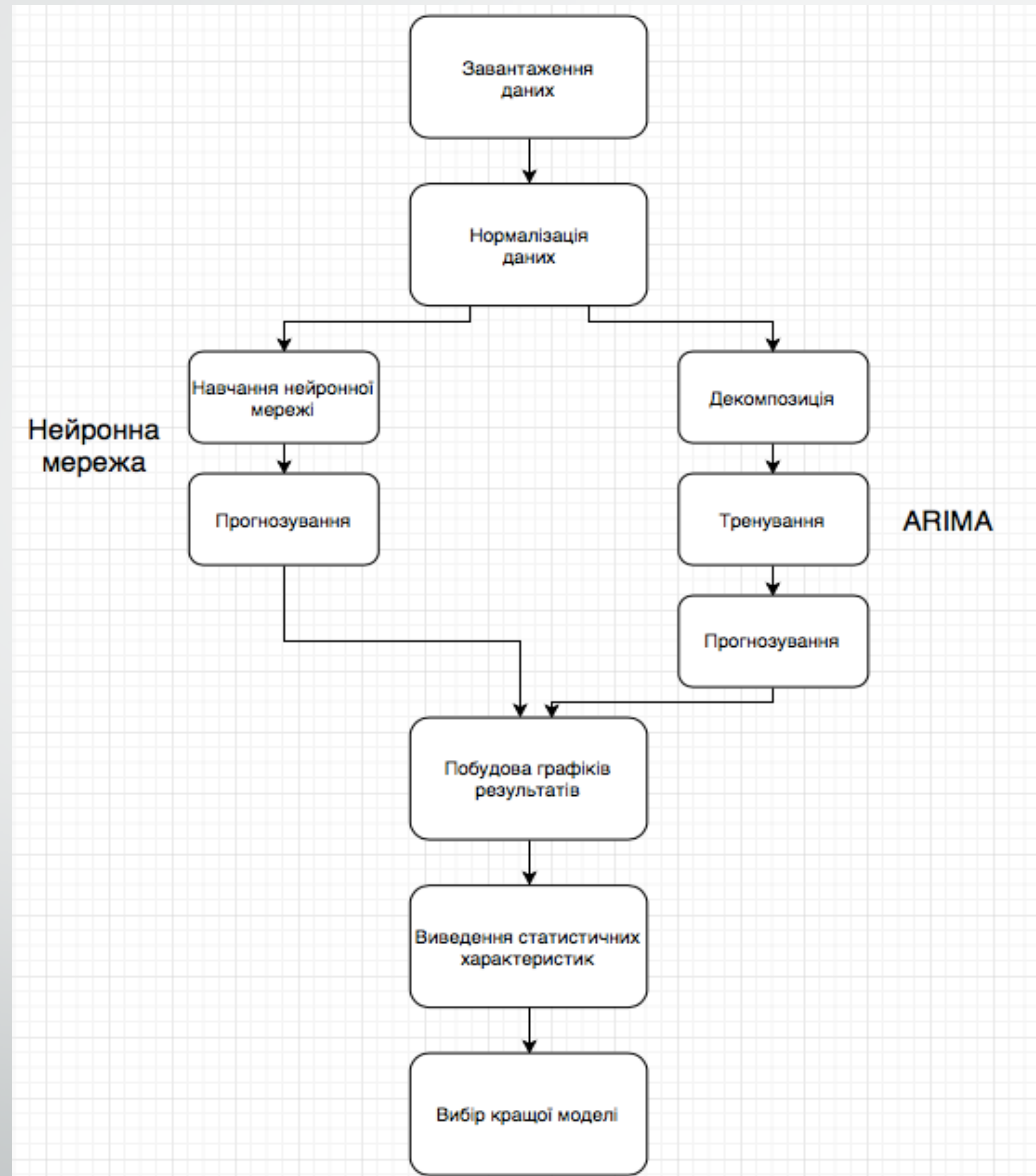
Про компанію Ryanair

- Ryanair є найбільшою в Європі бюджетною авіакомпанією, яка має штаб-квартиру в Ірландії. Ця компанія була заснована в 1985 році та починала свою діяльність з запуску щоденних польотів на 15-місному турбогвинтовому літаку між аеропортом Уотерфорд та Лондонським Гатвіком. В даний час вона є найбільшою недорогою авіакомпанією в Європі, яка базується на стратегії моделі Low Cost Leadership



Блок-схема процесу

5



Штучні нейронні мережі

- Нейронні мережі побудовані на простій моделі нейрона, що називається формальним нейроном.

Вихід формального нейрона розраховується за формулою:

$$y = f\left(\sum_{j=0}^{n-1} w_{i,j}x_j - \Theta_i\right)$$

де y - це вихід, $f(x)$ - нелінійна функція n - кількість вхідних ліній i —ого нейрона, а

Θ_i - поріг.

- Багатошаровий персептрон містить 3 типи нейронів: Вхідні, Вихідні та Приховані.

Алгоритм зворотного поширення помилки:

Спочатку вхідні дані поширюються по нейронній мережі до виходу.

$$e_k = d_k - y_k$$

де y_k - це розрахунковий результат, а d_k - це очікуваний вихід нейрона k

$$\delta_k = e_k g'(y_k)$$

де $g'(y_k)$ - це похідна функції активації.

$$\delta_j = \eta g'(y_j) \sum_{k=0}^K \delta_k \omega_{jk}$$

k - це кількість нейронів на даному шарі, η - це параметр швидкості навчання, який вказує, як потрібно скоригувати вагу.

$$\Delta\omega_{jk} = \delta_j y_k \text{ де } \Delta\omega_{jk} - \text{параметр коригування ваги } \omega_{jk} = \omega_{jk} + \Delta\omega_{jk}$$

Модель авторегресії з ковзним середнім

- складається з двох частин: авторегресивної (AR) частини та частини ковзного середнього (MA). Модель AR може бути написана наступним чином.

$$y_t = b_0 + b_1 y_{t-1} + b_2 y_{t-2} + \dots + b_p y_{t-p} + \varepsilon_t$$

- А модель MA може бути розписана як:

$$y_t = a_0 - a_1 \varepsilon_{t-1} - a_2 \varepsilon_{t-2} - \dots - a_q \varepsilon_{t-q} + \varepsilon_t$$

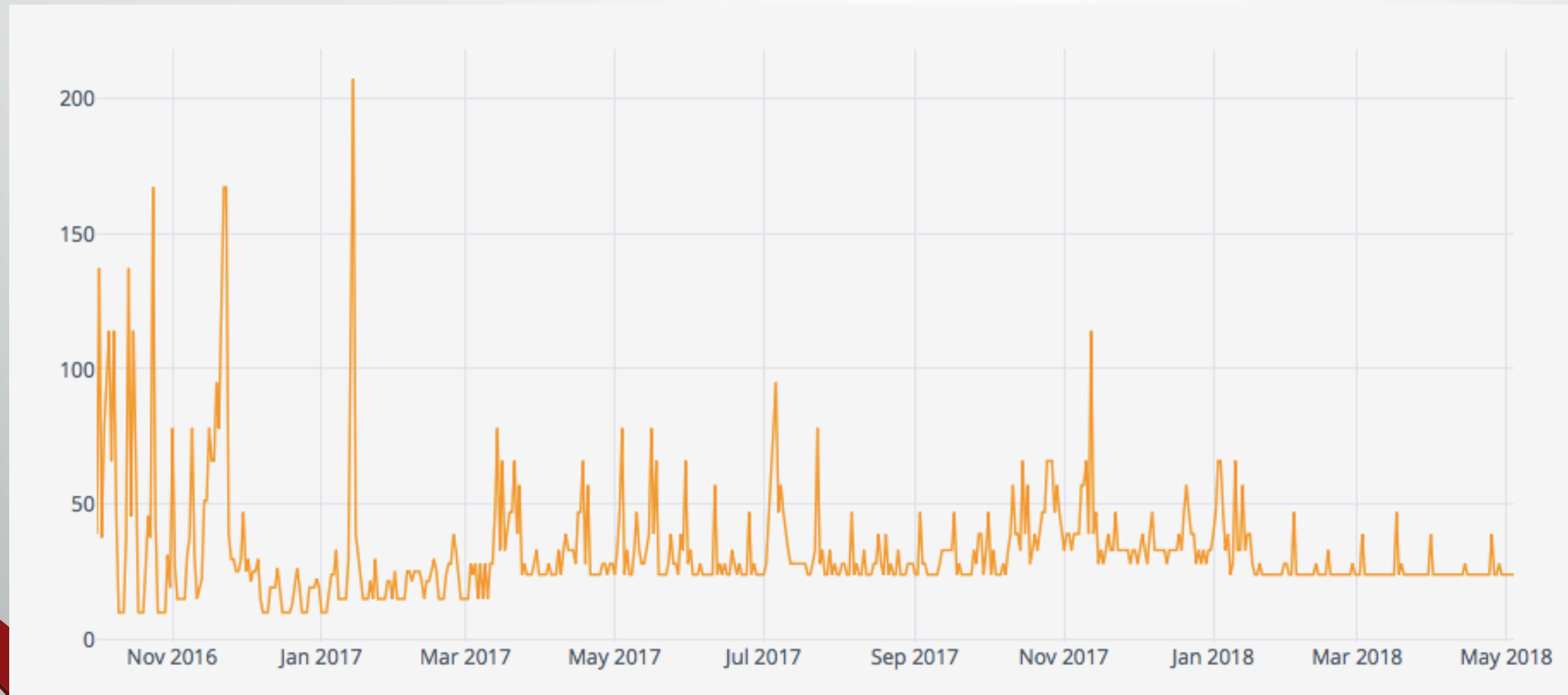
де y_t - значення часового ряду, $y_{t-1}, y_{t-2} \dots$ - попередні значення часового ряду
 $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}$ - залишкові члени a_0, a_1, a_2 - константи та коефіцієнти MA моделі
 b_0, b_1, b_2 - константи та коефіцієнти AR моделі.

- Математично модель ARIMA (p, q, d) представлена як

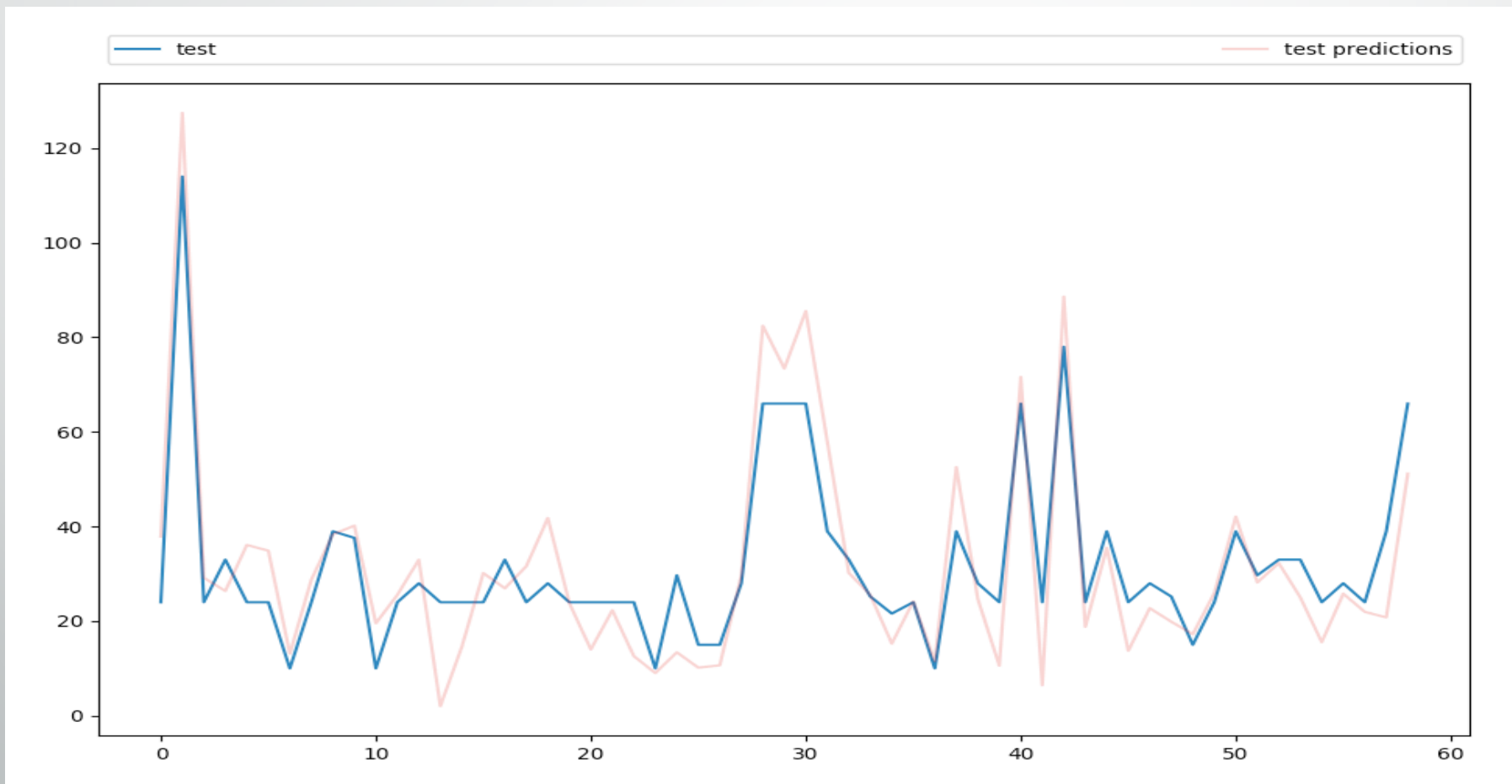
$$\Delta^d y_t = b_0 + \varepsilon_t + \sum_{i=1}^p b_i \Delta^d y_{t-i} + \sum_{j=1}^q a_j \varepsilon_{t-j}$$

p - це порядок авторегресивного процесу, q - це порядок процесу переміщення середнього, а d - процес диференціювання.

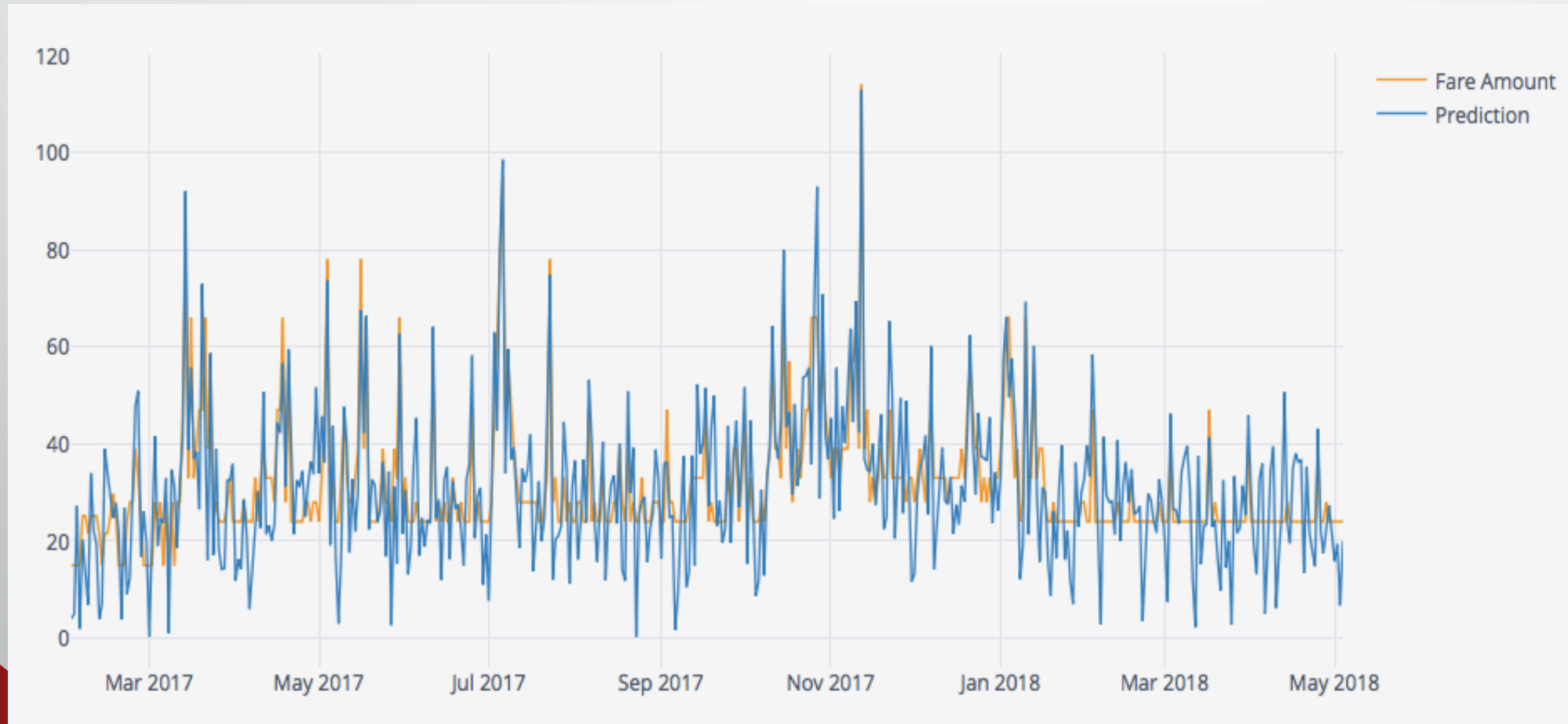
Вхідні дані (рейс Лондон-Мілан з жовтня 2016 по травень 2018 року)



Результуюча модель прогнозування за допомогою нейронних мереж



Результуюча модель прогнозування за допомогою ARIMA



Статистичні характеристики

ARIMA	BIC	AIC
(1, 1, 1)	4980.324	4958.614
(0, 1, 0)	5405.185	5396.5
(1, 1, 0)	5189.566	5172.198
(0, 1, 1)	5025.738	5008.370
(0, 0, 1)	3678.245	3624.332
(1, 0, 0)	2573.443	2491.702
(1, 0, 1)	1234,118	1171,29

ANN		MAE			MSE	
	100 epochs	500 epochs	1000 epochs	100 epochs	500 epochs	1000 epochs
10-10-1	2.93210	2.04163	3.12363	5.27054	5.93539	6.12363
10-14-1	1.93028	1.74002	1.82014	3.91674	3.37236	4.85576
10-17-1	1.62381	1.33832	1.57009	3.17878	2.400592	2.91639
10-20-1	1.74900	1.85531	2.210036	2.81792	2.93576	3.09010

Модель	R ²	MSE	MAE
ARIMA (1, 0, 1)	0.273863	3.340012	1.746113
ANN (10-17-1)	0.602384	2.400592	1.33832

Аналіз результатів

- Аналіз результатів показав, що як модель ARIMA, так і модель ШНМ можуть досягти хорошого прогнозування в застосуванні до реальних проблем і, таким чином, можуть ефективно залучати вигідне прогнозування цін на квитки. З графіків помітно, що розроблена прогнозована модель цін на авіаквитки з підходом на базі моделі ШНМ демонструвала вищу продуктивність, ніж модель ARIMA. Так як нейронні мережі вміють навчатися, що є одним з головних переваг перед іншими алгоритмами, такий результат був очікуваним.

Висновки

- Проведено порівняльний аналіз методів прогнозування ARIMA та моделі Штучних Нейронних Мереж
- Створено програмний продукт, який буде давати прогноз цін на авіаквитки
- Розроблена система, що дозволяє прогнозувати ціни на авіаквитки

Опубліковані тези:

- Чечельницька О.Б. Прогнозирование цен на авиабилеты бюджетной авиакомпании ruynaair с помощью нейронных сетей и авторегрессии со скользящим средним // Системний аналіз та інформаційні технології: Матеріали 20 Міжнар. наук.-техн. конф. SAIT 2018, 21-24 трав. 2018р. / М-во освіти і науки України [та ін.]. — К.: ННК «ІПСА» НТУУ «КПІ», 2018.

Шляхи подальшого розвитку

- 1. Використання інших методів прогнозування та їх порівняльна характеристика.
- 2. Вдосконалення моделей для досягнення більшої точності за рахунок зміни вибірки, інших методів нормалізації, кодування даних тощо.

Дякую за увагу!