

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ «КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ  
ІНСТИТУТ» ІМЕНІ ІГОРЯ СІКОРСЬКОГО  
«ІНСТИТУТ ПРИКЛАДНОГО СИСТЕМНОГО АНАЛІЗУ»

# **Інформаційна система аналізу тональності новин на основі системного підходу**

Автор: Рись Артем, група КА-61м  
Науковий керівник: д.т.н., професор Данилов В.Я.

**Об'єкт дослідження** - семантична орієнтація текстів новин.

**Предмет дослідження** - метод Наївного Байєсу та згорткові нейронні мережі.

**Актуальність обраної теми** - зростаюча популярність соціальних мереж, таких як Twitter, в останні роки викликало великий інтерес до аналізу настроїв їх користувачів. Аналіз тональності відноситься до класу методів, заснованих на обчислювальній обробці, що використовується для ідентифікації, вилучення, та характеристики суб'єктивної інформації, наприклад, думок, виражених в тій чи іншій частині тексту, що дозволяє класифікувати ставлення автора до різних тем в позитивні та негативні категорії. Аналіз настроїв має широку сферу застосувань в різних галузях, включаючи бізнес-аналітику, політику, соціологію тощо. Такі дані, як коментарі, блоги, твіти, відео і т.д., які висловлюють думки відносно різних тем і подій, пропонують величезні можливості для вивчення і аналізу людських думок і настроїв.

**Методи дослідження** - нейронні мережі, методи обробки текстів.

# Збір та обробка даних

2 набори даних:

- основний - фрагменти новин зі статей та заголовки новин статей про економіку США
- допоміжний - твіти на загальну тематику з Twitter

# Методи дослідження

## Наївний Байєс

$$p(C | F_1, \dots, F_n) = \frac{1}{Z} p(C) \prod_{i=1}^n p(F_i | C)$$

де

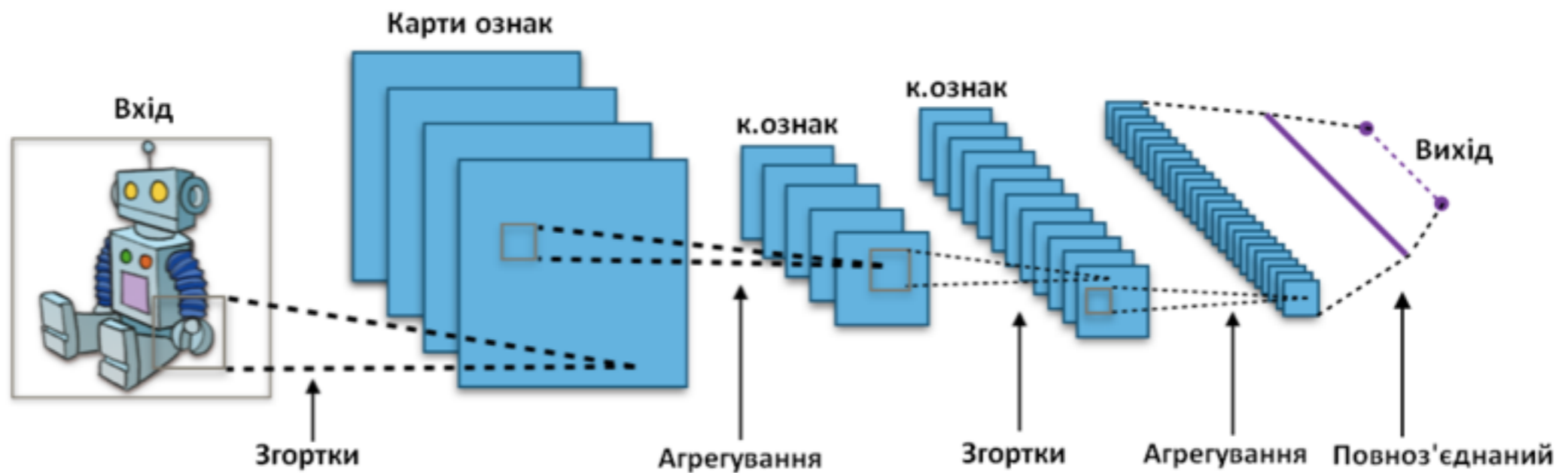
$Z$  - масштабуючий множник, який є константою,

$p(C)$  – ймовірність класу,

$P(F_i | C)$  – умовна ймовірність ознаки, за умови класу  $C$ .

# Методи дослідження

## Згорткові нейронні мережі



# Аналіз результатів

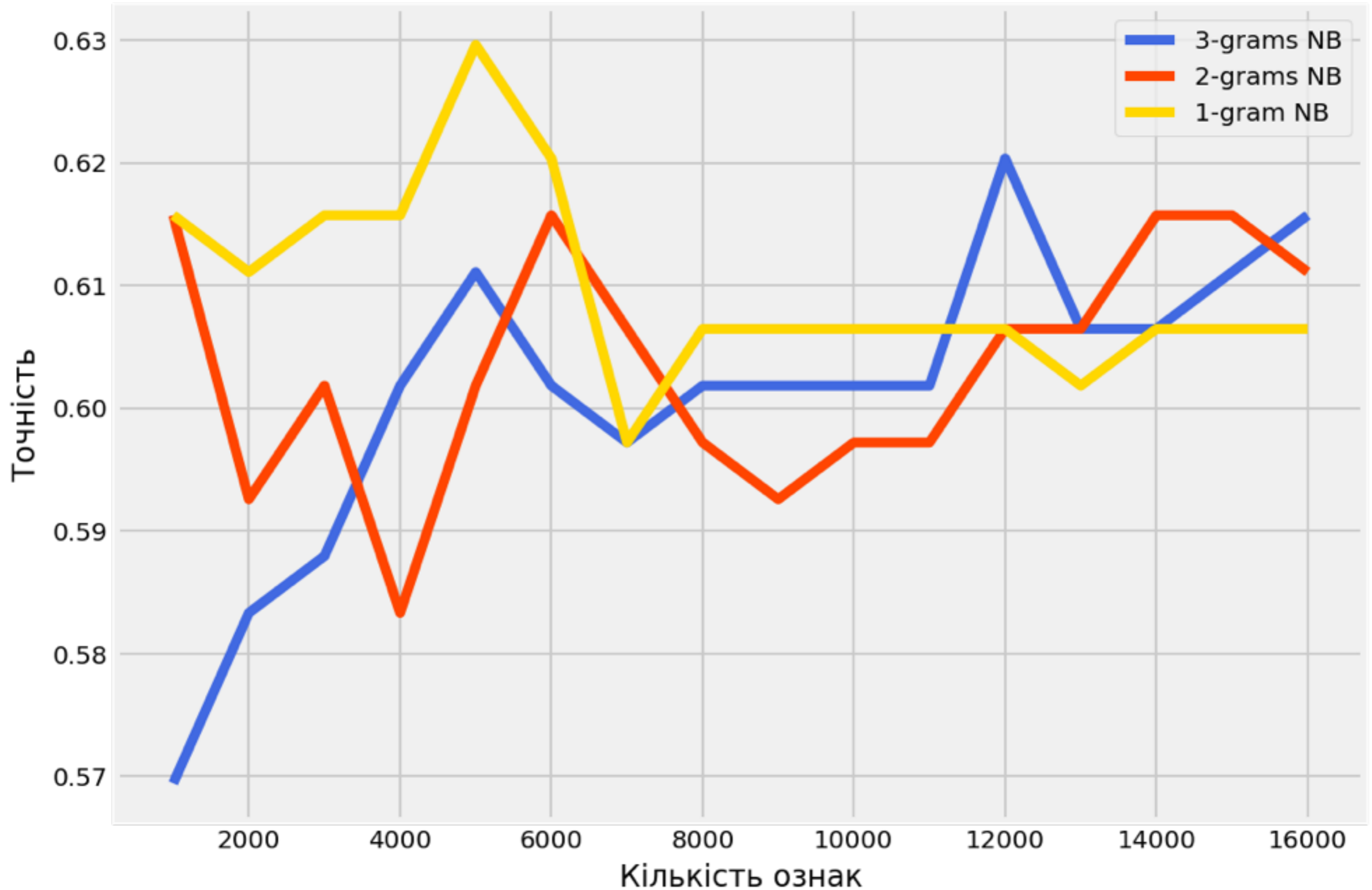
Метрики аналізу результатів:

- Точність
- Чутливість
- Специфічність
- F1 метрика

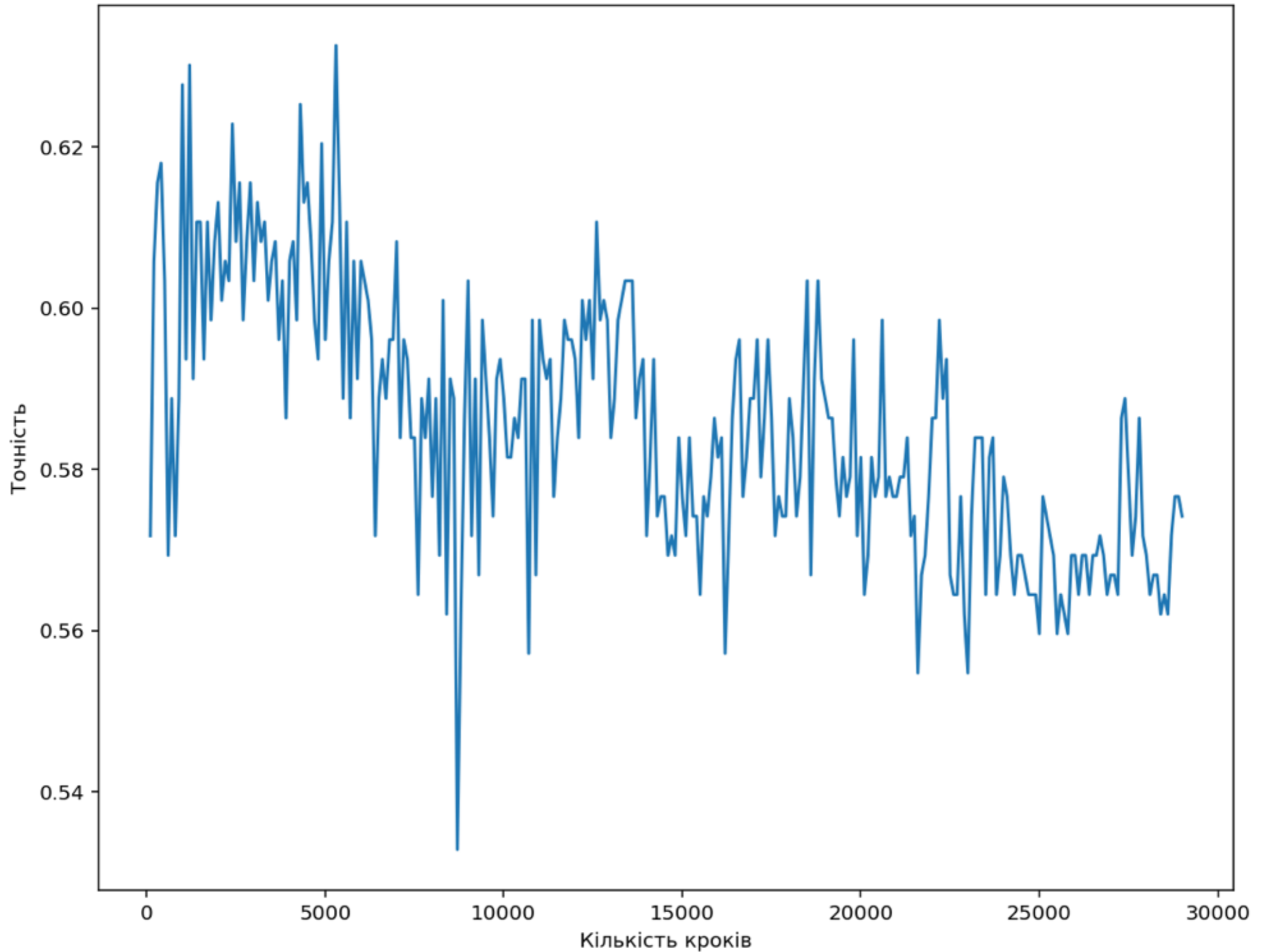
Набори даних:

- Новини
- Твіти

# Наївний Байєс

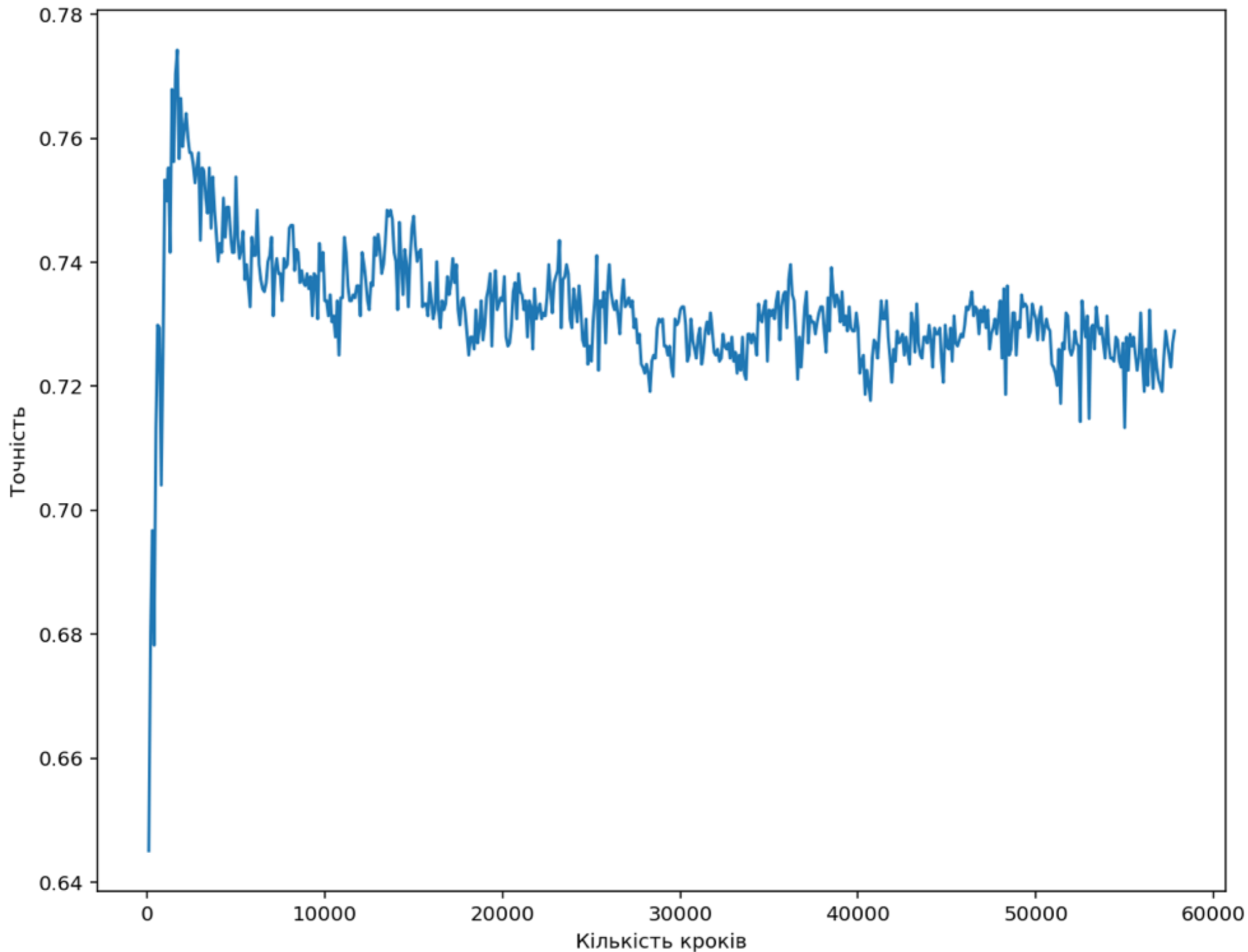


# Згорткові нейронні мережі (новини)





# Згорткові нейронні мережі (твіти)



# Перелік посилань

- 1.Sentiment analysis and opinion mining. Synthesis lectures on human language technologies [Електронний ресурс] / Liu B. — 2012. — № 5(1). — pp. 1-167. — Режим доступу: <https://www.morganclaypool.com/doi/pdf/10.2200/S00416ED1V01Y201204HLT016>.
- 2.Medhat W. Sentiment analysis algorithms and applications: A survey / Medhat W., Hassan A., Korashy H. // Ain Shams Engineering Journal. — 2014. — № 5(4). — pp. 1093-1113.
- 3.Turney P. D. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of review / Turney P. D. // Association for Computational Linguistics. — 2002. — № 5. — pp. 417-424.
- 4.Tang B. Toward optimal feature selection in naive Bayes for text categorization / Tang B., Kay S., He H. // IEEE Transactions on Knowledge and Data Engineering. — 2016. — № 28(9). — pp.2508-2521.
- 5.Kim Y. Convolutional neural networks for sentence classification / Kim Y. arXiv — 2014. — 20 p. — (Препринт / arXiv 2014-4).
- 6.Machine learning and lexicon based methods for sentiment classification: A survey / Hailong Z., Wenyan G., Bo J. — Washington: In Web Information System and Application Conference (WISA), 2014 11th. IEEE. — 2014. — pp. 262-265.
- 7.Mining and summarizing customer reviews / Hu M., Liu B. — In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM. — NY., 2004. — pp. 168-177.

**Дякую за увагу!**