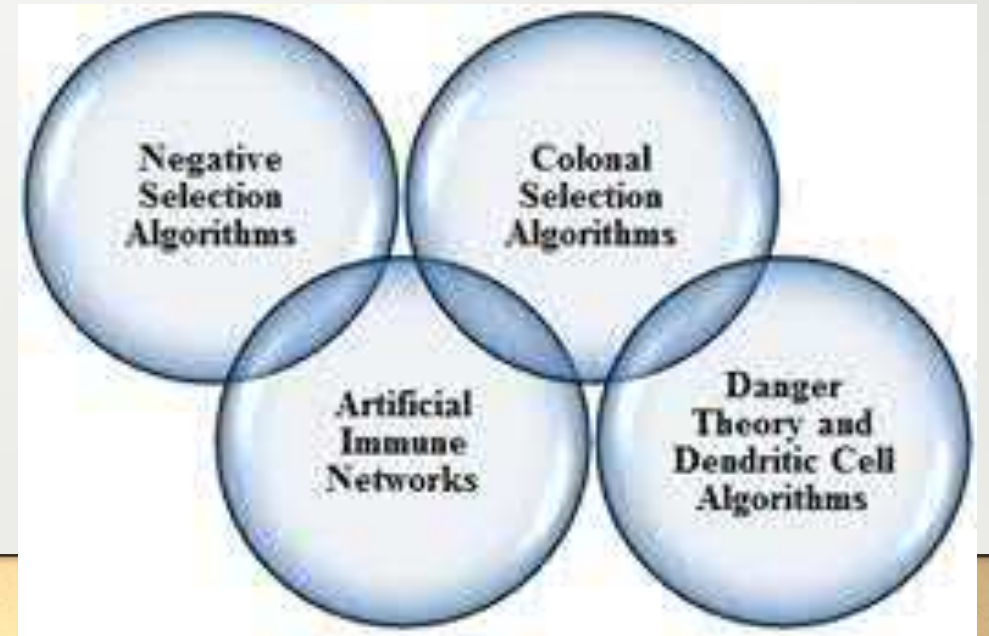




# МОДЕЛІ ТА МЕТОДИ АНАЛІЗУ КОРОТКИХ ВИБІРОК ШТУЧНИМИ ІМУННИМИ СИСТЕМАМИ

БАБІЧ МАРІЯ, КА-43

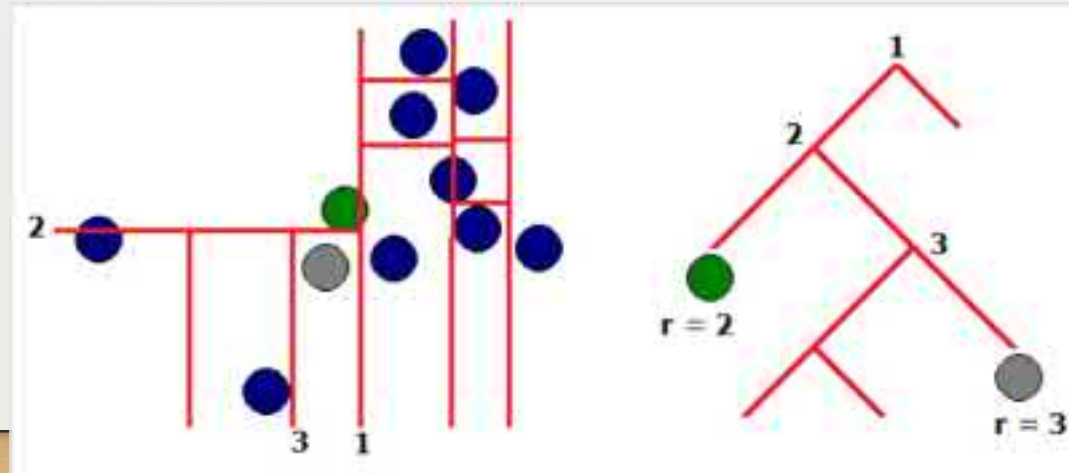
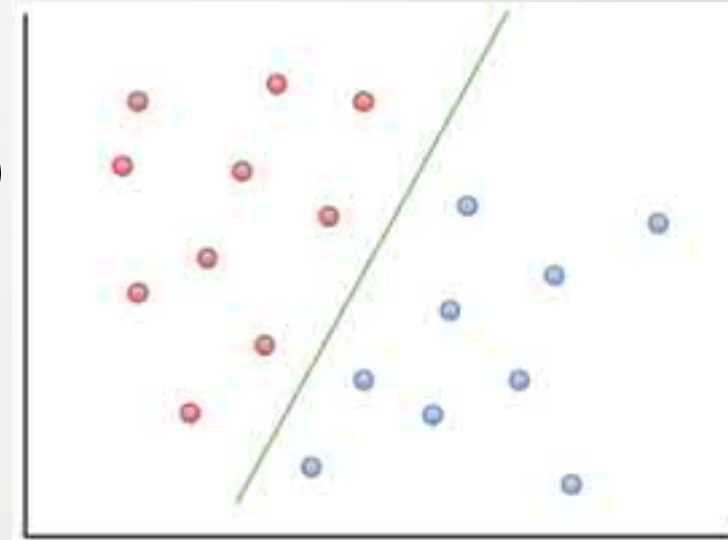
- **Актуальність** теми дослідження визначається впровадженням комп'ютерних технологій і систем штучного інтелекту в різні сфери діяльності людини. Ці дослідження спираються на фундаментальні основи сучасної науки такі, як розпізнавання образів, нейронні мережі, генетичні алгоритми, штучні імунні системи.



- **Предметом** дослідження є методи виявлення аномалій на незбалансованих вибірках в часових рядах, а саме дійснозначний метод негативної селекції.
- **Об'єктом** дослідження є вибірка, записами якої є показники мережі, записи серед яких є як нормальні, так і різномантні вторгнення та атаки.

# Існуючі Методи для розв'язання даної проблеми

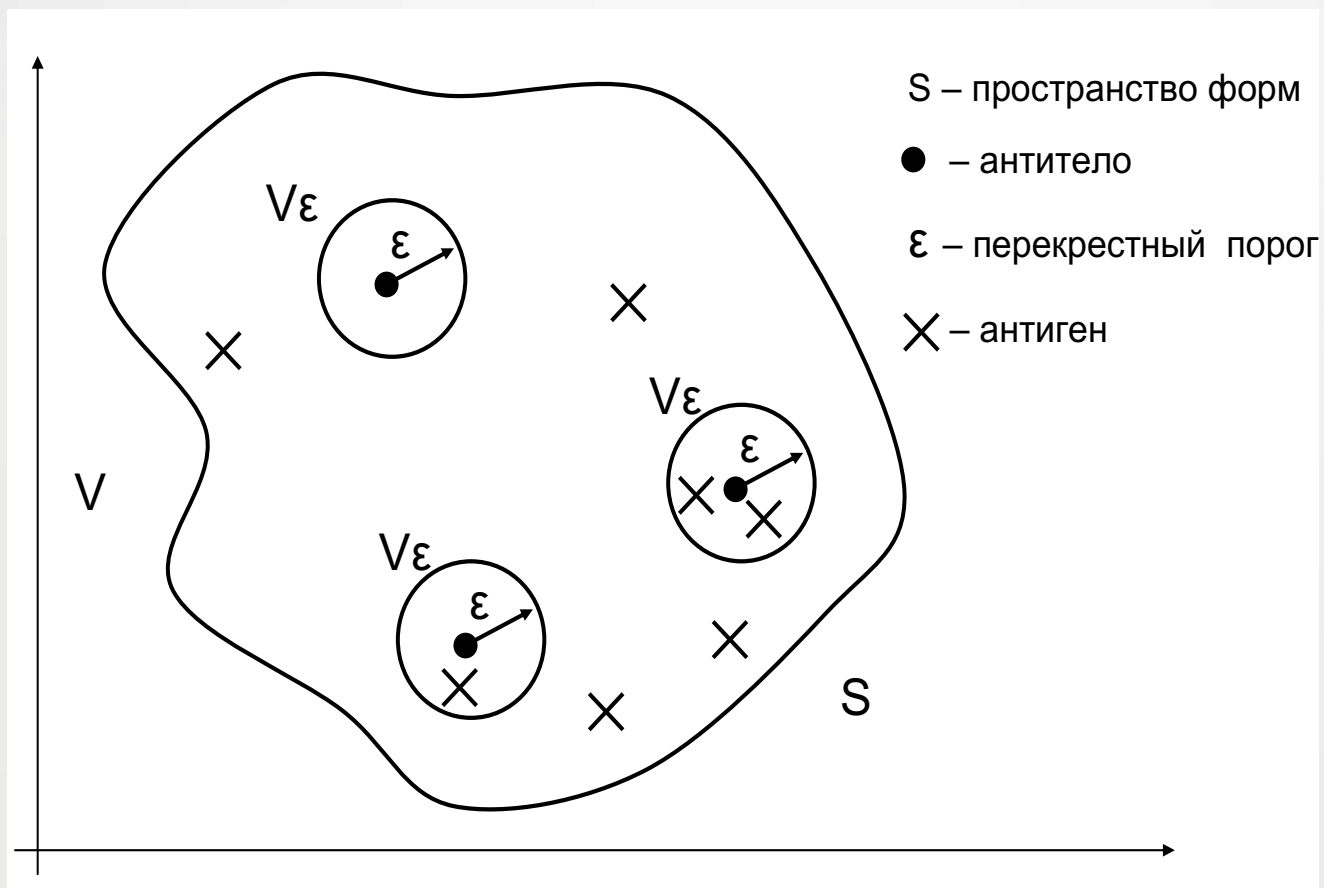
- Метод опорних векторів для одного класу (OneClassSVM)
- Ізолюючий ліс (IsolationForest)
- Еліпсоїдальної апроксимація даних (EllipticEnvelope)



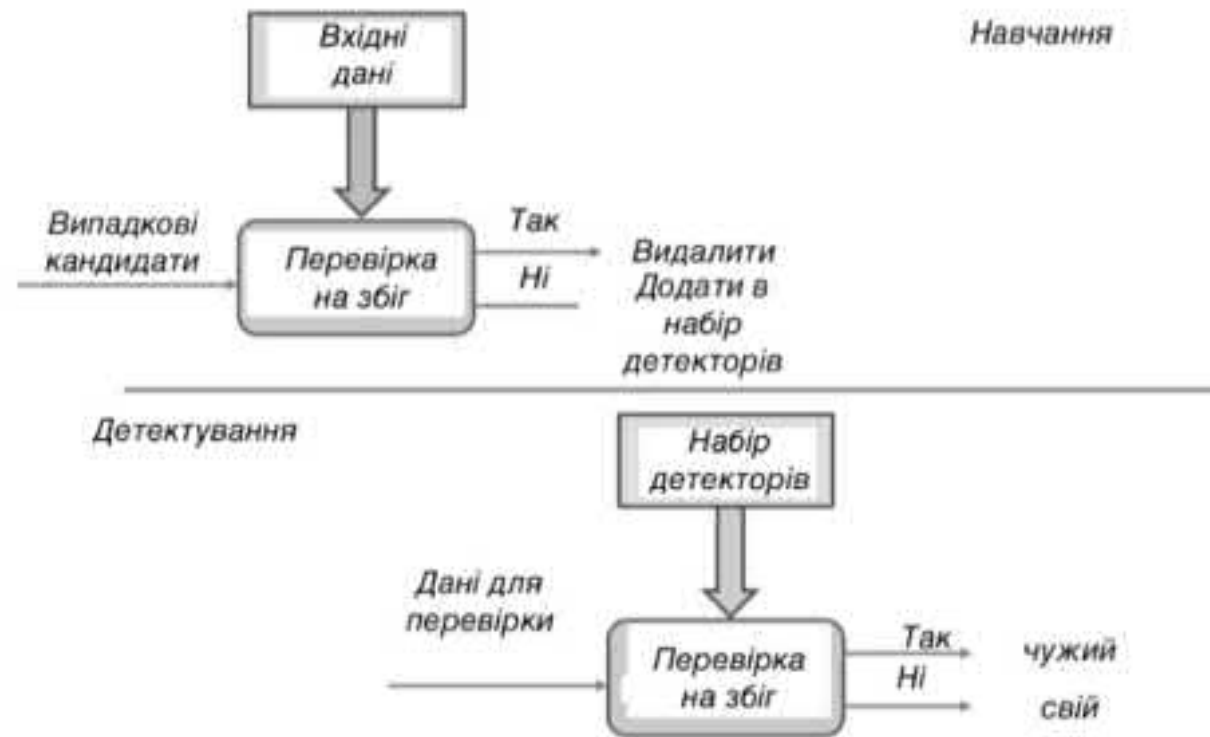
# ШТУЧНІ ІМУННІ СИСТЕМИ



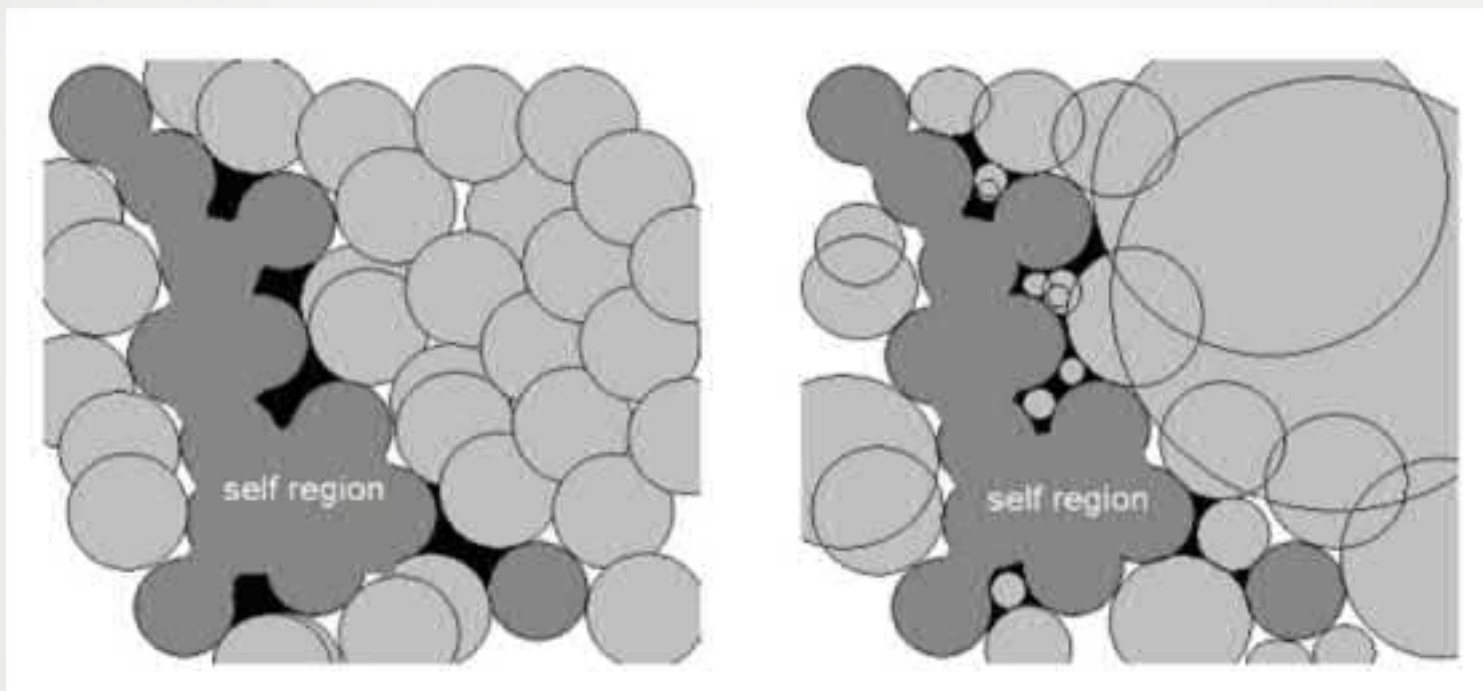
# Взаємодія антитіл та антигенів в просторі форм



# Алгоритм негативної селекції



Порівняння покриття простору детекторами з однаковими детекторами і детекторами з змінним радіусом





## ПЕРЕВАГИ МЕТОДУ

- Однокласова класифікація. Хоча цілю являється розділення на 2 класи, реально для навчання системи доступні лише зразки одного класу ('свій').
- Можливо фіксувати будь-які, навіть невідомі аномалії



# Недоліки методу

- 1. Негативний відбір буває неефективним і неминучі помилково позитивні результати.
- 2. Межі "свій" - "чужий" нечіткі, тому що найчастіше "своє" і "чуже" поділяють спільні області.
- 3. "Своє" з часом змінюється, тому очікуються проблеми з клітинами пам'яті, які з часом стануть неефективними.
- 4. Необхідно багато часу для тренування алгоритму, що в режимі реального часу або вимагає більші обчислювальні потужності, або змушує обирати інший алгоритм.

# Вибірка

- Вибірка була взята з <http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html>.
- Цей датасет використовувався для The Third International Knowledge Discovery and Data Mining Tools Competition. В даних інформація про деякі показники мережі, потрібно виявити вторгнення в мережу або певну атаку.



	0	1	2	3	4	5	6	7	8	9	...	32	33	34	35	36	37	38	39	40	41
0	0	tcp	http	SF	181	5450	0	0	0	0	...	9	1.0	0.0	0.11	0.0	0.0	0.0	0.0	0.0	normal.
1	0	tcp	http	SF	239	486	0	0	0	0	...	19	1.0	0.0	0.05	0.0	0.0	0.0	0.0	0.0	normal.
2	0	tcp	http	SF	235	1337	0	0	0	0	...	29	1.0	0.0	0.03	0.0	0.0	0.0	0.0	0.0	normal.
3	0	tcp	http	SF	219	1337	0	0	0	0	...	39	1.0	0.0	0.03	0.0	0.0	0.0	0.0	0.0	normal.
4	0	tcp	http	SF	217	2032	0	0	0	0	...	49	1.0	0.0	0.02	0.0	0.0	0.0	0.0	0.0	normal.


# Результати

	Метод негативної селекції	SVM
Час	Wall time: 10min 32s	Wall time: 6min 45s
Точність	Accuracy: 0.943279020559083	Accuracy: 0.932579020559083

## Висновки

- В роботі було розглянуто алгоритм штучних імунних систем, а саме дійснозначний алгоритм негативної селекції. Програмно реалізовано алгоритм дійснозначної негативної селекції.
- Результат порівняли з іншим методом (опорних векторних машин) і, на жаль, він дав кращі результати, оскільки потребує менше часу для навчання, а точність суттєво не відрізняється. Проте при наявності достатньої кількості обчислювальних потужностей, цей недолік втрачається.

- На даний момент штучні імунні системи досить нове поле для досліджень, і імунні алгоритми вивчаються та з'являються різноманітні модифікації, які є більш ефективними і мають більш вузьку спеціалізацію на тип даних і постановку задачі, яку треба вирішити.
- Оскільки метод негативної селекції – це метод класифікації, то в нас повинні бути помічені дані для навчання та тестування. Тому в майбутньому цікаво буде також розглянути інший алгоритм штучних імунних систем, а саме AiNet.



ДЯКУЮ ЗА УВАГУ!