

MLOps
ML
Pipeline

Vision &
Scope

Failure
Modes

Stress Testing
System Classification

Data Validation
Feature Selection

Model Compression
Explainability

Model Governance
Explainability

Reliability
Deployment

Weibull Distribution
Model Governance

Risk Assessment

КАФЕДРА КОМП'ЮТЕРНИХ НАУК



Марія Талах
Валентина Дворжак
Юрій Добровольський
Катерина Газдюк

ТЕХНОЛОГІЇ РОЗРОБКИ, ТЕСТУВАННЯ ТА ЕКСПЛУАТАЦІЇ ІНТЕЛЕКТУАЛЬНИХ ІНФОРМАЦІЙНИХ СИСТЕМ: ВІД ТЕОРІЇ НАДІЙНОСТІ ДО ML-РІШЕНЬ

Навчально-методичний посібник



МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧЕРНІВЕЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ЮРІЯ ФЕДЬКОВИЧА

М.В. Талах, В.В. Дворжак,
Ю.Г. Добровольський, К.П. Газдюк

ТЕХНОЛОГІЇ РОЗРОБКИ,
ТЕСТУВАННЯ ТА
ЕКСПЛУАТАЦІЇ
ІНТЕЛЕКТУАЛЬНИХ
ІНФОРМАЦІЙНИХ СИСТЕМ:
ВІД ТЕОРІЇ НАДІЙНОСТІ ДО
ML-РІШЕНЬ

Навчальний посібник

Технодрук
2025

УДК 004.8:004.9:681.3.06

Т-38

Рекомендовано до друку вченою радою Навчально-наукового інституту фізико-технічних та комп'ютерних наук Чернівецького національного університету імені Юрія Федьковича (протокол №17 від 18 грудня 2025 р.)

Рецензенти:

Виклюк Я.І., доктор технічних наук, професор, професор кафедри систем штучного інтелекту Національного університету “Львівська політехніка”

Баловсяк С. В., доктор технічних наук, доцент, доцент кафедри комп'ютерних систем та мереж Чернівецького національного університету імені Юрія Федьковича

Талах М.В., Дворжак В.В., Добровольський Ю.Г., Газдюк К.П.,

Т-38 Технології розробки, тестування та експлуатації інтелектуальних інформаційних систем: від теорії надійності до ML-рішень. Навчальний посібник/ Талах М.В., Дворжак В.В., Добровольський Ю.Г., Газдюк К.П., – Чернівці: Чернівецький нац.ун-т, 2025. – 411 с.

/

М.В. Талах, В.В. Дворжак, Ю.Г. Добровольський,
К.П. Газдюк – Чернівці: Технодруk, 2025. – 411 с.

Навчальний посібник присвячений дослідженню фундаментальних аспектів технологій розробки, тестування та експлуатації інтелектуальних інформаційних систем. Він охоплює як теоретичні засади надійності, так і практичні підходи до створення, комплексного тестування та підтримки сучасних ML-рішень.

Навчальний посібник призначений для студентів, які навчаються за напрямками підготовки "Комп'ютерні науки" та "Інженерія програмного забезпечення", а також для всіх, хто цікавиться розробкою, впровадженням та підтримкою надійних систем машинного навчання.

ЗМІСТ

<u>ВСТУП</u>	<u>11</u>
<u>РОЗДІЛ 1. ФУНДАМЕНТАЛЬНІ ПОНЯТТЯ І ВИЗНАЧЕННЯ ТЕОРІЇ НАДІЙНОСТІ В КОНТЕКСТІ ML-СИСТЕМ</u>	<u>16</u>
1.1. Визначення поняття "надійність" та її особливості для інтелектуальних систем.....	16
1.2. Класифікація і характеристики відмов у традиційному ПЗ та ML-системах	21
1.3. Термінологія теорії надійності.	23
1.4. Бізнес-аналіз та визначення вимог до надійності ML-систем	25
1.4.1. Візія та межі проекту (Vision and Scope)	25
1.4.2. Бізнес-вимоги та їх вплив на надійність	25
1.4.3. Трансформація бізнес-цілей у цілі машинного навчання	30
1.4.4 Ключові патерни застосування машинного навчання	33
1.5. Методологія CRISP-ML(Q) та її роль у життєвому циклі ML-систем.....	37
<u>РОЗДІЛ 2. КЛАСИФІКАЦІЯ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ ТА ЇХ НАДІЙНІСТЬ</u>	<u>41</u>
2.1. Проблема побудови систематики технічних систем	41

2.2. Класифікація технічних систем за функцією та принципом дії.....	46
2.3. Класифікація технічних систем за рівнем складності	47
2.4. Класифікація технічних систем за способом виготовлення.....	49
2.5. Класифікація технічних систем за ступенем оригінальності	49
2.6. Класифікація технічних систем за типом виробництва	50
2.7. Класифікація ПЗ за його місцем у технічному процесі	51
2.8. Класифікація інтелектуальних технічних систем (ML/AI).....	52
2.9. Аналіз здійсненості проекту на основі Machine Learning Canvas framework	57

РОЗДІЛ 3. ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ ТА ЙОГО КЛАСИФІКАЦІЯ.....61

3.1. Класифікація програмного забезпечення інтелектуальних систем	61
3.2. Системне програмне забезпечення для ML-систем	61
3.3. Прикладне програмне забезпечення та його роль в ML-системах	64
3.4. Системи програмування для ML-розробки..	66
3.5. Створення РОС та бачення рішення (vision of the solution).....	67
3.5.1. Vision Statement	67

3.5.2. ML and Non-ML Components in a System	69
3.5.3. ML Pipelines describing	73
3.5.4. Основні системні компоненти	84
3.5.5. Основні функціональні і нефункціональні можливості продукту	87
3.5.6. Припущення та Залежності (Assumptions and Dependencies)	93
3.5.7. Обсяг і обмеження (scope and limitations)	93
3.6. Вимоги користувача до надійних ML-систем	96
3.6.1. Опис типових сценаріїв використання системи	96
3.6.2. Створення беклогу продукту	100

РОЗДІЛ 4. КРИТЕРІЇ НАДІЙНОСТІ ТА ЇХ ЗАСТОСУВАННЯ ДО ML-СИСТЕМ106

4.1. Критерії надійності невідновних систем	107
4.2. Критерії надійності відновлюваних об'єктів	110
4.3. Вибір даних (Data selection)	112
4.3.1. Збір та верифікація даних	112
4.3.2. Аналіз даних та опис вимог до даних	113
4.3.3. Документування статистичних властивостей даних	114
4.4. Оцінка якості даних та їх впливу на надійність ML-системи	115

РОЗДІЛ 5. ЗАКОНИ РОЗПОДІЛУ ЧАСУ ДО ВІДМОВИ В КОНТЕКСТІ ML-СИСТЕМ122

5.1. Життєвий цикл технічного засобу та ML-системи	122
5.2. Розподіл Вейбулла та його застосування в ML-системах.....	123
5.3. Експоненційний розподіл для оцінки надійності ML-моделей	125
5.4. Розподіл Релея в контексті ML-систем	127
5.5. Гамма-розподіл та його роль в оцінці надійності.....	128
5.6. Нормальний розподіл для моделювання відмов ML-систем.....	130
5.7. Трикутний розподіл	131
5.8. Закони розподілу дискретних випадкових величин.....	132

РОЗДІЛ 6. ІНЖЕНЕРІЯ ДАНИХ (DATA ENGINEERING) ДЛЯ НАДІЙНИХ ML-СИСТЕМ 135

6.1. Feature extraction та його вплив на надійність	135
6.2. Feature selection та оптимізація надійності .	137
6.3. Балансування класів (Class balancing).....	143
6.4. Очищення даних	146
6.5. Інженерія ознак (Feature engineering)	153
6.6. Аугментація даних (Data augmentation).....	158

РОЗДІЛ 7. ТЕСТУВАННЯ ТА НАДІЙНІСТЬ ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ІНТЕЛЕКТУАЛЬНИХ СИСТЕМ 165

7.1. Основні поняття та фактори надійності ML-систем	165
7.2. Модель аналізу надійності програмних засобів ML-систем.....	166
7.3. Фактори, що впливають на надійність ПЗ та ML-систем.....	167
7.4. Помилки ПЗ та їх прояви в ML-системах ...	167
7.5. Засоби та способи підвищення надійності ПЗ ML-систем.....	171
7.6. Проблеми дослідження надійності ПЗ ML-систем	172
7.7. Тестування ПЗ ML-систем	172
7.8. Фреймворки для тестування ML-систем	174
7.9. Класифікація показників якості ПЗ ML-систем	181
7.10. Якість моделей машинного навчання та показники якості моделей	183
7.11. Якість даних та показники якості даних...	186
7.12. Тестування конвеєрів машинного навчання та показники якості конвеєрів	188
7.13. Основні показники якості системи з ML-компонентами та системне тестування.....	191
7.14. Основні показники надійності ПЗ ML-систем	195
7.15. Загальна характеристика та класифікація моделей надійності ПЗ ML-систем	198
7.16. Моделі надійності з урахуванням недосконалого відлагодження ML-систем.....	205
7.17. Моделі надійності з урахуванням розподілу зусиль тестування ML-систем	208

РОЗДІЛ 8. МОДЕЛІ НАДІЙНОСТІ ПЗ ДЛЯ ML-СИСТЕМ.....214

8.1. Динамічні моделі надійності та їх застосування в ML-системах..... 215

8.2. Статичні моделі надійності для ML-систем 218

РОЗДІЛ 9. РОЗРОБКА МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ З УРАХУВАННЯМ НАДІЙНОСТІ223

9.1. Визначення якісних показників моделі.....223

9.2. Вибір ML-алгоритмів (baseline selection)225

9.3. Додавання доменних знань для спеціалізації моделі229

9.4. Навчання моделі з урахуванням вимог надійності.....238

9.5. Застосування transfer learning242

9.6. Компресія моделі.....247

9.7. Ансамблеве навчання.....253

РОЗДІЛ 10. ОЦІНЮВАННЯ МОДЕЛЕЙ МАШИННОГО НАВЧАННЯ З ТОЧКИ ЗОРУ НАДІЙНОСТІ257

10.1. Валідація продуктивності моделі.....257

10.2. Визначення стійкості ML-моделі260

10.3 Підвищення пояснюваності моделі263

10.4. Прийняття рішення щодо розгортання моделі266

10.5. Документування навчання та оцінювання моделі.....	268
---	-----

РОЗДІЛ 11. РОЗГОРТАННЯ НАДІЙНИХ ML-СИСТЕМ (DEPLOYMENT).....270

11.1. Визначення вимог до апаратного забезпечення для інференції.....	270
11.2. Оцінка моделі в продуктивному середовищі (онлайн-тестування)	272
11.3. User Acceptance and Usability Testing.....	273
11.4. Розробка плану резервного копіювання.....	274
11.5. Налаштування стратегії розгортання.....	274
11.6. Інтеграція в робочі процеси MLOps.....	275
11.7. Оцінка моделі у виробничих умовах	275
11.8. Забезпечення прийняття користувачами та зручності використання	284
11.9. Управління моделлю (Model governance)..	295
11.10. Розгортання за обраною стратегією	305
11.11 DevOps практики для ML-систем (MLOps)	319

РОЗДІЛ 12. МОНІТОРИНГ ТА ПІДТРИМКА НАДІЙНОСТІ ML-СИСТЕМ322

12.1. Моніторинг ефективності та результативності передбачень моделі	323
12.2. Порівняння з критеріями успіху	340
12.3 Технології моніторингу та обслуговування ML-систем у промисловому середовищі	353

12.4. Перенавчання моделі за необхідності.....	354
12.5. Збір нових даних.....	359
12.6. Маркування нових даних	363
12.7. Повторення завдань із фаз "Model Engineering" та "Model Evaluation"	371
12.8. Безперервна інтеграція, навчання та розгортання моделі	374
12.9. Технології та інструменти CI/CD для ML-систем	379
<u>ВИСНОВКИ</u>	<u>390</u>
<u>ДОДАТКИ</u>	<u>396</u>
Додаток 1. Контрольні запитання.....	396
Додаток 2. Приклади тестових завдань	399
<u>СПИСОК ЛІТЕРАТУРИ</u>	<u>410</u>