

О.О. Матвієнко, доц., канд. техн. наук, **А.В. Гриньків**, ст. досл., канд. техн. наук,
О.М. Лівіцький, канд. техн. наук

Центральноукраїнський національний технічний університет, м. Кропивницький, Україна
e-mail: richdad.ua@gmail.com

Методика вибору робастних моделей вібродіагностики за доменного зсуву та шуму міток

У статті розглянуто проблему зниження якості моделей машинного навчання для вібродіагностики вузлів мобільних машин за умов доменного зсуву (зміна режимів роботи, навантаження, стендових/польових умов, датчиків) та шуму міток (неточні записи технічного сервісу, помилки класифікації станів, перехідні режими). Запропоновано умовну методику вибору архітектури моделі у вигляді матриці рішень, яка пов'язує тип і “суворість” доменного зсуву, обсяг і якість розмічених даних, рівень зашумлення сигналу та міток із доцільним класом моделей (класичні моделі на інженерних ознаках, гібридні підходи, ансамблеві та напівнаглядкові стратегії). Методика підтверджена експериментами на задачах діагностики підшипників і редукторів: для помірного доменного зсуву показано перевагу робастифікованих ансамблів за високого шуму міток, тоді як для сильного доменного зсуву встановлено обмежену ефективність складних ансамблевих стратегій і доцільність простіших рішень із доменною адаптацією. Отримані результати формують практичні рекомендації для побудови інтелектуальної системи технічного сервісу, орієнтованої на обслуговування за станом у реальних польових умовах.

мобільні машини, вібродіагностика, доменний зсув, шум міток, робастність, ансамблеві моделі, напівнаглядкове навчання, псевдорозмітка, підшипники кочення, редуктор, інтелектуальна система технічного сервісу

Постановка проблеми. Підвищення надійності мобільних машин підприємств агропромислового виробництва потребує переходу від регламентного та реактивного обслуговування до обслуговування за технічним станом [1]. У таких умовах ключову роль відіграють дані сенсорної діагностики (зокрема, вібраційні сигнали вузлів обертання), за якими методи машинного навчання можуть виявляти перехідні та несправні стани. Однак практичне впровадження моделей діагностики у польових умовах ускладнюється двома системними чинниками: доменним зсувом та шумом міток.

У цій роботі терміни «робастна модель» і «робастність» використовуються у загальноприйнятому для технічних і прикладних наук значенні: це здатність моделі зберігати прийнятну якість за малих або помірних змін умов (режимів роботи, параметрів знімання сигналу) та за наявності помилок/неточностей у розмітці. Відповідно «робастифікація» — це цілеспрямоване підвищення робастності моделі (наприклад, через двоступеневе навчання з псевдорозміткою та фільтрацією за впевненістю).

Доменний зсув виникає через зміну режимів роботи, навантажень, швидкостей, типів датчиків і способів монтажу, а також через відмінності між стендовими та експлуатаційними умовами. Внаслідок цього статистичні властивості ознак і сигналів у цільовому домені відрізняються від навчального, що призводить до деградації якості моделі при перенесенні. Другим чинником є шум міток, обумовлений неповнотою або неточністю записів технічного сервісу, помилками класифікації станів, а також наявністю перехідних режимів, що погано піддаються однозначній розмітці. За таких умов «універсальний» вибір моделі (наприклад, фіксований вибір глибокої мережі) є недостатнім: доцільність архітектури повинна визначатися умовно, залежно від характеристик доменного зсуву та якості даних.

Аналіз останніх досліджень і публікацій. У попередніх роботах [2–3] було окреслено, по-перше, які саме діагностичні сигнали доцільно використовувати для оцінки технічного стану мобільних машин та вузлів (вібраційні, акустичні, температурні, струмові тощо), а по-друге — як організувати дані (збір, передавання, зберігання, «паспорт» набору даних, версійність, відтворюваність) так, щоб вони реально підтримували інтелектуальний технічний сервіс і машинне навчання. У роботі [2] наголошено на необхідності переходу від порогових підходів до інтелектуальних алгоритмів і системи технічного сервісу, що інтегрує сенсорну базу та моделі для онлайн-оцінки стану. У роботі [3] підкреслено практичну проблему: профільні набори даних часто є фрагментарними й слабко задокументованими, тому потрібна узгоджена «доріжка» від польового збору до аналітики з контролем якості, метаданими та відтворюваними релізами даних.

У літературі з діагностики роторних машин і підшипників важливим є висновок про підвищення інформативності за рахунок поєднання каналів спостереження: вібрації, акустичної емісії та аналізу продуктів зносу в мастилi [4]. Водночас найбільш поширеною практикою у прикладних роботах залишається схема “вібрація → ознаки → класифікація стану”, де інформацію про дефект витягують через часові та спектральні характеристики [5]. Оглядові роботи з вібродіагностики підкреслюють, що методики моніторингу стану зазвичай добре працюють у контрольованих умовах, однак їх результативність істотно залежить від режимів роботи, навантаження та особливостей вимірювальної постановки (тип датчика, монтаж, параметри знімання сигналу) [6]. Окремо зростає інтерес до застосування методів машинного навчання в акустико-емісійних вимірюваннях як до додаткового (або альтернативного) джерела діагностичної інформації [7].

Загалом можна виділити два “пласти” підходів. Перший — класичний ML на інженерних ознаках, де головну роль відіграють якісне виділення інформативних характеристик у часовій, частотній і часо-частотній областях та їх відбір (зменшення розмірності, відсів зайвих ознак). Другий — глибоке навчання, яке прагне автоматизувати побудову ознак із сирих сигналів, але є більш вимогливим до репрезентативності даних і ресурсів. Для задач промислової діагностики нерідко застосовують також гібридні інженерні рішення (наприклад, поєднання спектральних уявлень з нечіткою логікою), що дозволяють формалізувати евристичні експертів [8].

Показовим прикладом “прагматичного” підходу є робота [9], де запропоновано схему діагностики на базі LightGBM із багатодоменим виділенням ознак та ітеративним відбором (модифікований RFE), а також спеціальною валідацією за режимами навантаження (leave-one-loading-out), що підсилює здатність моделі узагальнювати на нові, не представлені під час навчання умови. У цій роботі автори прямо фіксують проблему: при зміні режиму між навчанням і тестом якість різко падає, а контроль відбору ознак і валідації за режимами дозволяє зробити модель стійкішою.

Окремою лінією досліджень є діагностика за змінних умов роботи (швидкість, навантаження, інші датчики/стенди), що в сучасних термінах описують як доменний зсув (domain shift). Суть проблеми проста: модель навчають на «джерельних» даних (source domain), а застосовують на «цільових» (target domain), і ці дані вже не з одного розподілу. У роботі [10] ця ситуація описана як порушення базового припущення “train і test з одного розподілу”, що особливо проявляється при зміні робочих умов. Як вихід пропонується доменна адаптація, зокрема безнаглядне узгодження доменів через вирівнювання підпросторів (subspace alignment), щоб класифікатор, навчений на source, працював на target без повної перенавчальної розмітки.

Близькими за ідеологією є підходи domain-adversarial transfer learning, де модель навчається так, щоб ознаки були “схожими” між доменами. У роботі [11] доменна

адаптація розглядається як спосіб перенесення діагностичних моделей між машинами/умовами з економією розмітки. Автори вводять у контекст діагностики Domain-Adversarial Neural Networks (DANN) та наголошують, що без таких механізмів моделі схильні перенавчатися й деградувати на нових машинах і режимах. Також запропоновано уніфікований протокол експериментальної перевірки для порівняння методів адаптації.

Лінію доменної адаптації розвивають роботи зі спеціалізованими варіантами адаптації для підшипників: domain-conditioned adaptation для складних режимів [12], dynamic domain adaptation для змінних умов [13], multi-domain adversarial transfer із урахуванням інформації про механізми відмов [14], а також комбіновані схеми з одночасним використанням CNN, метрик розбіжності розподілів і domain-adversarial навчання [15]. Такі підходи важливі тим, що вони підсилюють переносимість моделей, але водночас підвищують вимоги до даних: потрібні узгоджені протоколи експериментів, відслідковування доменів і метаданих, а також контроль якості розмітки та релізів даних.

Другий критичний чинник практики — шум міток (label noise). На відміну від шуму сигналу (перешкоди датчика), шум міток означає, що частина прикладів у навчальному наборі має неправильний клас: помилки в журналах сервісу, «перехідні» стани, суб'єктивність експертної розмітки, змішані режими. Через це модель може вчитися на суперечливих прикладах і втратити узагальнюваність. У роботі [16] прямо зазначено, що в реальних застосуваннях дані часто є дисбалансними (мало прикладів рідкісних відмов) і водночас сильно зашумленими, через що традиційні схеми діагностики демонструють низьку якість. В якості відповідей на ці виклики розглядаються покращення підготовки даних, стратегії балансування, а також робастні схеми навчання з використанням синтетичних даних і гібридних моделей.

Поряд із повною класифікацією станів у літературі розвиваються й підходи виявлення аномалій (anomaly detection), де модель навчається переважно на «нормі», а відхилення трактуються як потенційні несправності. Для акустичних і вібраційних даних промислових процесів показано ефективність автоенкодерів, у яких аномалії проявляються через зростання похибки реконструкції [17]. Для транспортних застосувань розглядаються й сценарії раннього виявлення несправностей за звуковими сигналами автомобіля [18]. Окремо підкреслюється важливість явного врахування доменного зсуву при оцінюванні систем виявлення аномалій у машинному моніторингу (зокрема, у постановці DCASE 2021 Task 2 навчальні та тестові дані належать до різних доменів) [19]. Також показано, що синтетичне розширення вибірки може підвищувати стійкість підходів до нових умов та варіацій запису [20]. Як приклад модальності, що добре масштабуються на транспортні задачі, наведено й підхід аудіодіагностики двигуна з підбором ознак та оптимізованими класифікаторами [21].

Для мобільних машин агропромислового виробництва зазначені проблеми часто посилюються організаційними обмеженнями: сезонність, різноманітність умов, обмежена сенсорна оснащеність, розподіленість даних між власником техніки, сервісом і виробником. У роботі [22] звертається увага, що в агросекторі дані про машини часто розпорошені, а сенсорні дані можуть бути недостатні для сервісних провайдерів; як альтернатива пропонується організація прогностичного обслуговування на основі сервісних записів і платформи взаємодії «фермер–сервіс».

Таким чином, огляд показує: існує значний спектр моделей — від ознакового машинного навчання (зокрема моделей градієнтного бустингу) до глибоких нейронних мереж; активно розвиваються методи перенесення і адаптації між режимами та стендам; окремо розробляються підходи до дисбалансу й шуму розмітки та до виявлення аномалій. Проте в багатьох публікаціях ці напрями подані як «окремі

рішення”, і рідше — як інженерна методика вибору, що підказує: коли достатньо простішої моделі з контролем якості даних, а коли без доменної адаптації і робастифікації вже не обійтися. Саме цей розрив між “наявністю методів” і “наявністю зрозумілих правил вибору під реальні обмеження” і формує мотивацію нашої роботи, де доменний зсув і шум міток розглядаються спільно в контексті інтелектуального технічного сервісу.

Постановка завдання. Метою роботи є розроблення умовної методики (правил) вибору робастних моделей вібродіагностики вузлів мобільних машин за наявності доменного зсуву та шуму міток, придатної для практичного впровадження в інтелектуальній системі технічного сервісу [1–3].

Для досягнення поставленої мети необхідно розв’язати такі завдання:

1. Формалізувати основні фактори, що визначають складність діагностичної задачі в польових умовах: тип/рівень доменного зсуву (зміна режимів, навантажень, датчиків, стенд–експлуатація), обсяг і якість розмічених даних, рівень зашумлення сигналів і міток;

2. Визначити класи моделей і стратегій навчання, доцільні за різних комбінацій зазначених факторів: моделі на інженерних ознаках (зокрема, моделі градієнтного бустингу), гібридні рішення, робастні ансамблеві підходи, напівнаглядові схеми, а також методи доменної адаптації/перенесення [9–15];

3. Побудувати матрицю рішень як інженерний інструмент вибору моделі та процедури навчання з урахуванням умов недоцільності застосування складних стратегій (дефіцит даних, відсутність контролю якості, ресурсні обмеження);

4. Перевірити працездатність методики на задачах діагностики підшипників кочення та редукторів із контрольованим варіюванням доменного зсуву й шуму міток, порівнявши базові та робастифіковані підходи за узгодженими метриками якості;

5. Сформулювати практичні рекомендації щодо інтеграції методики в контур технічного сервісу (зокрема, вимоги до даних, метаданих і процедур контролю якості, необхідних для відтворюваного застосування матриці рішень) [3].

Об’єктом дослідження є процес інтелектуальної оцінки технічного стану вузлів мобільних машин за вібраційними сигналами. Предметом дослідження є методи та правила вибору робастних моделей діагностики за умов доменного зсуву і шуму міток.

Виклад основного матеріалу. Моніторинг та обслуговування за станом (CBM – condition-based maintenance) у цій роботі розглядаються як практична основа підвищення надійності мобільних машин: рішення про обслуговування приймається не лише за напрацюванням, а за фактичними ознаками деградації вузлів, що визначаються за сенсорними сигналами та моделями діагностики [1–3]. Нижче подано підхід, який дозволяє “прив’язати” вибір моделі та процедури навчання до реальних умов збору даних — насамперед до доменного зсуву та шуму міток.

Нехай маємо джерельний домен S (source domain) — набір даних, на якому навчається модель, та цільовий домен T (target domain) — дані, на яких модель застосовується. Доменний зсув означає, що розподіл даних у S і T різняться: змінюються режими роботи, навантаження, швидкості, датчики, умови монтажу, або ж відрізняються стендові та експлуатаційні умови. Наслідок цього — зниження якості при перенесенні моделі, навіть якщо в S вона показує високу точність.

У прикладній діагностиці доцільно розрізняти принаймні два “типи” зсуву: зсув за режимом/умовами (зміна навантаження, частоти обертання, режиму редуктора або стенду) і зсув за ступенем деградації, коли джерельних даних представлено одну “глибину” дефекту, а в цільових – іншу.

За інженерним змістом ці два типи впливають на модель по-різному: при зсуві за режимом/умовами змінюється “фон” і співвідношення компонент у сигналі, тоді як при

зсуві за ступенем деградації змінюється власне прояв дефекту. Саме тому в літературі активно розвиваються методи доменної адаптації та перенесення, які намагаються узгодити S і T без повної повторної розмітки [10–15].

Шум міток трактуємо як ситуацію, коли частина прикладів у навчальному наборі має неправильний клас. На практиці це наслідок помилок у сервісних журналах, неоднозначності “перехідних” режимів, людського фактору, а також зміни критеріїв віднесення стану до класу. Важливо відрізнити шум міток від шуму сигналу (перешкоди датчика): перший “ламає” навчальну ціль, тоді як другий знижує інформативність ознак. У поєднанні з доменним зсувом шум міток робить небезпечним пряме застосування складних моделей без перевірки їхньої стійкості, що особливо актуально для польових умов [3,16].

Виходячи з практики прикладної вібродіагностики та огляду літератури, розглянуто наступні базові сімейства рішень.

1. Класична модель на інженерних ознаках: формування набору часових і спектральних характеристик (статистичні моменти, енергії в діапазонах, спектральні індекси тощо) та навчання класифікатора типу градієнтного бустингу дерев (LightGBM). Цей підхід є відносно стійким до помірного шуму міток і дає відтворюваний базовий рівень якості за помірних ресурсів; додатковою перевагою є можливість контролю/відбору ознак та інтерпретації їхнього внеску [9].

2. Ансамблеві/розширені ознаки (naive ensemble): збільшення розмірності простору ознак (наприклад, за рахунок поєднання груп ознак або їхніх взаємодій) і навчання більш “ємної” моделі. Практичний ризик такого підходу — підвищена чутливість до помилок розмітки та до несумісності масштабів ознак, тому нормалізація/стандартизація стає обов’язковою умовою (це прямо узгоджується з вимогами до відтворюваного пайплайна, сформульованими у [3]).

3. Двоступенева робастифікація з псевдорозміткою: на першому кроці базова “стійка” модель (зазвичай на інженерних ознаках) формує псевдомітки для даних, де розмітка ненадійна або відсутня. На другому кроці навчається більш “ємна” модель (ансамбль/гібрид), але вже на псевдомітках, які додатково контролюються за рівнем впевненості. Така схема є простим інженерним компромісом між робастністю (1-й крок) і здатністю моделі витягувати складні закономірності (2-й крок), але її ефективність залежить від якості псевдоміток.

4. Доменна адаптація для ознак: коли зсув домену значний, навіть “стійкі” класичні моделі можуть різко деградувати. У такому разі доцільно застосовувати прості та відтворювані засоби узгодження доменів (наприклад, CORAL для вирівнювання коваріацій), а вже після цього — навчання моделі на узгоджених ознаках [10–15]. У нашій роботі цей елемент використано саме як інженерний «мінімум» для сильного доменного зсуву.

Запропонована методика реалізована як матриця рішень, що поєднує три групи умов:

- характер доменного зсуву (тип і «суворість»): зсув за режимом/умовами або за ступенем деградації; помірний чи сильний);
- стан розмітки (обсяг розмічених даних, очікуваний рівень шуму міток);
- обмеження впровадження (можливість валідації на цільових умовах, вимоги до відтворюваності пайплайна, ресурсні обмеження).

Логіка матриці має форму правил «якщо... то...» і включає також умови недоцільності застосування окремих стратегій. Узагальнено це можна подати так:

- якщо доменний зсув помірний, а в наявності є хоча б мінімальна можливість контролю якості (валідація/калібрування), то доцільно починати з класичної моделі на ознаках; при зростанні шуму міток переходити до двоступеневої робастифікації

(псевдорозмітка + фільтрація за впевненістю), оскільки «ємні» моделі більш чутливі до помилкових класів [9, 16].

– якщо доменний зсув сильний (наприклад, інший режим редуктора або інший стенд), то «переускладнення» архітектури без адаптації домену часто не дає вигаду. У такому разі базовим рішенням є класична модель + доменна адаптація ознак (мінімально інвазивний і відтворюваний крок), а робастифіковані ансамблі застосовуються лише як другий рівень — за наявності достатньої впевненості у псевдомітках.

– якщо базова модель у цільових умовах демонструє занадто низьку якість (або немає можливості її оцінити), псевдорозмітка здатна «закріпити» помилки й погіршити результат. У таких ситуаціях доцільніше зосередитись на збиранні невеликого «довіреного» набору міток у цільовому домені та/або на простіших, більш контрольованих рішеннях.

Таблиця 1 – Матриця рішень вибору моделей вібродіагностики за доменного зсуву та шуму міток

Доменний зсув	Доступні дані	Якість міток	Рекомендована модель/стратегія	Очікуване значення F1
Зсув навантаження (режиму)	Даних достатньо (>50%)	Мітки чисті	Гібридна модель або LightGBM	0.98 – 1.00
Зсув навантаження (режиму)	Даних мало (10–20%)	Мітки зашумлені ($\approx 10\%$)	Гібрид + псевдорозмітка	0.90 – 0.95
Зсув за ступенем дефекту	Даних достатньо (>50%)	Мітки чисті	Гібрид (обов'язково)	0.95 – 0.96
Зсув за ступенем дефекту	Даних достатньо (>50%)	Мітки чисті, але сигнал зашумлений	Гібрид (обов'язково)	0.40 – 0.50
Зсув за ступенем дефекту	Даних мало (10–20%)	Будь-яка	LightGBM (з обережністю)	< 0.70

Джерело: розроблено авторами

Перевірку методики виконано на двох сценаріях із контрольованим доменним зсувом та ін'єкцією шуму міток (симетричне випадкове спотворення частини навчальних міток). Як основну метрику використано Macro-F1, що є інформативною для багатокласової задачі та не маскує провали по окремих класах.

Сценарій А (помірний зсув): CWRU, зсув за ступенем дефекту.

У цьому сценарії використано публічний набір даних CWRU (Case Western Reserve University Bearing Data Center) [23]. Джерельний домен відповідав меншому дефекту підшипника, а цільовий — більшому (0.007" \rightarrow 0.021"). Тут 0.007" і 0.021" — це діаметр штучно сформованого дефекту (у дюймах), який у CWRU використовується для позначення ступеня пошкодження; таким чином моделюється помірний зсув за ступенем дефекту. Розглянуто три рівні шуму міток у навчанні: 15%, 20%, 25%. Порівнювалися: класична модель (LightGBM на інженерних ознаках), наївний ансамбль, робастифікований ансамбль із псевдорозміткою та робастифікований підхід із фільтрацією псевдоміток за впевненістю (поріг 0.85).

Таблиця 2 – Результати для CWRU за високого шуму міток (метрика Macro-F1)

Рівень шуму, %	LightGBM	Гібрид	Наївний ансамбль	Робастифікований ансамбль	Робастифікований ансамбль + фільтрація
15	0.6171	0.6609	0.7059	0.6902	0.6969
20	0.5750	0.8735	0.7032	0.8200	0.9368
25	0.6152	0.6003	0.6051	0.7667	0.8018

Джерело: розроблено авторами

Як видно з табл. 2, за шуму міток 20% застосування фільтрації псевдоміток дає найбільший ефект (приріст відносно наївного ансамблю становить близько 14% за Macro-F1), тобто фільтрація справді «очищує» навчальний сигнал для 2-го етапу робастифікації. Важливо, що класична модель у цьому сценарії демонструє відносно стабільний рівень якості, підтверджуючи її роль «надійного першого кроку» для генерації псевдоміток.

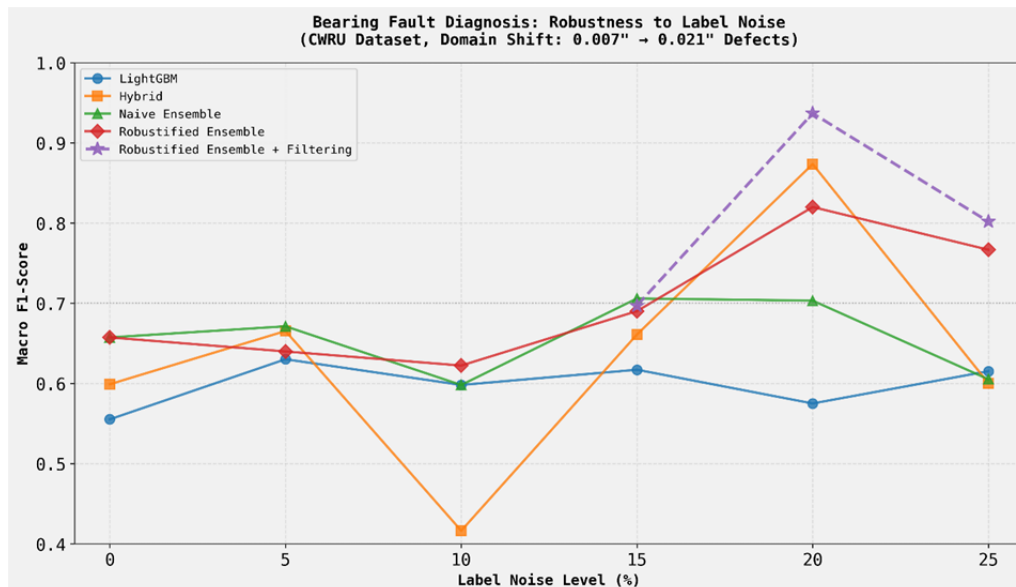


Рисунок 1 – Залежність Macro-F1 від рівня шуму міток для сценарію CWRU (помірний доменний зсув)
Джерело: розроблено авторами.

Сценарій В (сильний зсув): SEU, cross-condition (20 Hz–0 V → 30 Hz–2 V).

Сильний доменний зсув змодельовано на публічних даних редуктора SEU [24] при переході між двома режимами роботи (20 Hz–0 V → 30 Hz–2 V), де змінюються динаміка збудження та «фон» сигналу. Позначення 20–0 та 30–2 відповідають парі «швидкість обертання – налаштування навантаження» на стенді: 20 Hz і 30 Hz — режим швидкості, а 0 V і 2 V — керувальна напруга навантажувального вузла (що задає рівень навантаження/моменту), а не вихід вібродатчика. Для зниження розбіжності між доменами для інженерних ознак застосовано просту доменну адаптацію типу CORAL, після чого оцінено вплив шуму міток 15%, 20%, 25% для класичної моделі та двох ансамблевих стратегій.

Таблиця 3 – Результати для SEU за сильного доменного зсуву (Macro-F1)

Рівень шуму, %	Класична модель (CORAL+LightGBM)	Наївний ансамбль	Робастифікований ансамбль
15	0.5461	0.3453	0.3618
20	0.5435	0.3769	0.3651
25	0.5445	0.3491	0.4141

Для сильного доменного зсуву спостерігається інша картина, ніж у сценарії А: класична модель на адаптованих ознаках є найстабільнішою і майже не змінюється з ростом шуму міток, тоді як ансамблеві підходи суттєво поступаються їй. Двоступенева робастифікація дає нестійкий ефект: невелике покращення відносно наївного ансамблю при 15% і найбільше — при 25% шуму, але погіршення при 20% шуму. Це підтверджує практичний висновок матриці рішень: за сильного зсуву доцільно починати з простішої, контрольованої моделі з адаптацією домену, а складні ансамблеві стратегії застосовувати обережно.

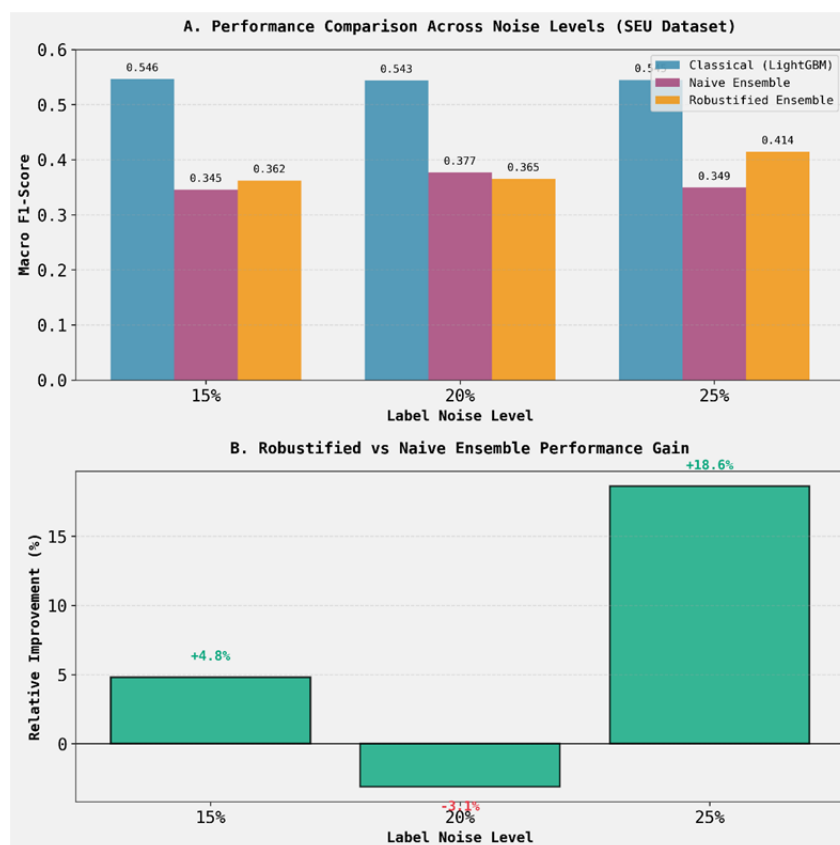


Рисунок 2 – Порівняння Macro-F1 для сценарію SEU за різного рівня шуму міток та відносний виграш робастифікації відносно наївного ансамблю

Джерело: розроблено авторами

Узагальнюючи обидва сценарії, можна зробити інженерний висновок: (1) за помірною зсуву робастифіковані ансамблі з контролем якості псевдоміток можуть давати суттєвий виграш при високому шумі розмітки; (2) за сильного зсуву «безпечнішим» вибором є класична модель на добре задокументованих інженерних ознаках із простою доменною адаптацією; (3) фільтрація за впевненістю підвищує стабільність двоступеневої робастифікації, але не скасовує потребу в коректно організованих даних і валідації на цільових умовах [3].

Висновки.

1. Обґрунтовано, що для вібродіагностики вузлів мобільних машин у реальних умовах доменний зсув і шум міток є не винятком, а типовими чинниками, які системно знижують якість моделей при перенесенні з навчального набору на експлуатаційні дані.

2. Запропоновано умовну методику вибору моделей у вигляді матриці рішень, що пов'язує «суворість» доменного зсуву, доступність і надійність розмітки, а також обмеження впровадження з доцільним класом рішень (моделі на інженерних ознаках, робастифіковані ансамблі, напівнаглядні стратегії, доменна адаптація).

3. Для сценарію CWRU з помірним доменним зсувом за ступенем дефекту показано, що робастифікований ансамбль із фільтрацією псевдоміток за впевненістю може суттєво підвищувати якість при високому шумі міток; максимальний ефект зафіксовано при 20% шуму (Macro-F1 = 0.937).

4. Для сценарію SEU зі сильним зсувом між режимами (cross-condition) встановлено, що найстабільнішим рішенням є класична модель на інженерних ознаках із простою доменною адаптацією (CORAL+LightGBM), тоді як ансамблеві стратегії демонструють нижчу якість і нестійкий ефект робастифікації.

5. Сформовано практичні рекомендації для інтеграції в інтелектуальну систему технічного сервісу: (а) починати з відтворюваного базового пайплайна та дисципліни даних/метаданих; (б) застосовувати псевдорозмітку лише за наявності можливості валідації на цільових умовах; (в) за сильного доменного зсуву пріоритетно використовувати узгодження доменів та контрольовані моделі, а «ускладнення» архітектури виконувати поетапно.

6. Подальші дослідження доцільно спрямувати на перевірку методики на польових даних мобільних машин та на розширення матриці рішень для інших модальностей (акустика, температура, струмові сигнали), а також на уточнення критеріїв кількісної оцінки «суворості» доменного зсуву.

Список літератури

1. Аулін В. В., Гриньків А. В., Головатий А. О., Лисенко С. В., Голуб Д. В., Кузик О.В., Тихий А. А. Методологічні основи проектування та функціонування інтелектуальних транспортних і виробничих систем: монографія під заг. ред. д.т.н., проф. Ауліна В.В. Кропивницький: Видавець Лисенко В.Ф., 2020. 428с.
2. Матвієнко О. О., Аулін В. В. Класифікація типів сигналів та методів машинного навчання для інтелектуальної оцінки технічного стану мобільних машин підприємств агропромислового виробництва. *Збірник наукових праць. Науковий вісник. Технічні науки*. 2025. № 11(42)_II. С. 298–312. DOI: 10.32515/2664-262X.2025.11(42).2.298-312.
3. Матвієнко О. О., Аулін В. В., Гриньків А.В. Стан та напрями розвитку архітектури даних для інтелектуальної оцінки технічного стану мобільних машин підприємств агропромислового виробництва. *Збірник наукових праць. Науковий вісник. Технічні науки*. 2025. № 12(43)_I. С. 227–237. DOI: 10.32515/2664-262X.2025.12(43).1.227-237.
4. Loutas T. H., Roulias D., Pauly E., Kostopoulos V. The combined use of vibration, acoustic emission and oil debris on-line monitoring towards a more effective condition monitoring of rotating machinery // *Mechanical Systems and Signal Processing*. 2011. Vol. 25, no. 4. P. 1339–1352. DOI: 10.1016/j.ymsp.2010.11.007.
5. Oyedaja K. O. Diagnostics of Bearing Defects Using Vibration Signal // *International Journal of Computer and Electrical Engineering*. 2012. Vol. 4, no. 6. P. 821–825. DOI: 10.7763/IJCEE.2012.V4.612.
6. Sen A., Majumder M. C., Mukhopadhyay S., Biswas R. K. Condition Monitoring of Rotating Equipment Considering the Cause and Effects of Vibration: A Brief Review // *International Journal of Modern Engineering Research (IJMER)*. 2017. Vol. 7, Iss. 1. P. 36–49.
7. Ciaburro G., Iannace G. Machine-learning-based methods for acoustic emission testing: A review. *Applied Sciences*. 2022. Vol. 12, No. 20. 10476. DOI: 10.3390/app122010476.
8. Da Silva R. R., Da S. Costa E., De Oliveira R. C. L., Mesquita A. L. A. Fault Diagnosis in Rotating Machine Using Full Spectrum of Vibration and Fuzzy Logic // *Journal of Engineering Science and Technology*. 2017. Vol. 12, no. 11. P. 2952–2964.
9. Saberi A. N., Belahcen A., Sobra J., Vaimann T. LightGBM-Based Fault Diagnosis of Rotating Machinery Under Changing Working Conditions Using Modified Recursive Feature Elimination // *IEEE Access*. 2022. Vol. 10. P. 81910–81925. DOI: 10.1109/ACCESS.2022.3195939.
10. Zhang B., Li W., Tong Z., Zhang M. Bearing fault diagnosis under varying working condition based on domain adaptation // *arXiv preprint*. 2017. arXiv:1707.09890.
11. Wang Q., Michau G., Fink O. Domain Adaptive Transfer Learning for Fault Diagnosis // *2019 Prognostics and System Health Management Conference (PHM-Paris)*. IEEE, 2019. DOI: 10.1109/PHM-Paris.2019.00054.
12. Zhang X., Gu G. Fault Diagnosis for Rolling Bearings Under Complex Working Conditions Based on Domain-Conditioned Adaptation // *Machines*. 2024. Vol. 12, no. 11. Art. 787. DOI: 10.3390/machines12110787.
13. Zhang Q., Lv Z., Hao C., Yan H., Fan Q. Intelligent Fault Diagnosis of Bearings in Unsupervised Dynamic Domain Adaptation Networks Under Variable Conditions // *IEEE Access*. 2024. Vol. 12. P. 82911–82925. DOI: 10.1109/ACCESS.2024.3413087.
14. Zhong X. Failure Mechanism Information-Assisted Multi-Domain Adversarial Transfer Fault Diagnosis Model for Rolling Bearings under Variable Operating Conditions // *Sensors*. 2024. Vol. 24, no. 3. Art. 1036. DOI: 10.3390/s24031036.
15. Li X., Wang J., Wang J., Wang J., Li Q., Yu X., Chen J. Research on Unsupervised Domain Adaptive Bearing Fault Diagnosis Method Based on Migration Learning Using MSACNN-IJMMD-DANN // *Machines*. 2025. Vol. 13. Art. 618. DOI: 10.3390/machines13070618.
16. Jalayer M., Kaboli A., Orsenigo C., Vercellis C. Fault Detection and Diagnosis with Imbalanced and Noisy Data: A Hybrid Framework for Rotating Machinery // *Machines*. 2022. Vol. 10. Art. 237. DOI: 10.3390/machines10040237.

17. Duman T. B., Bayram B., İnce G. Acoustic Anomaly Detection Using Convolutional Autoencoders in Industrial Processes // *Advances in Intelligent Systems and Computing*. 2020. Vol. 1028. P. 397–406. DOI: 10.1007/978-3-030-20055-8_41.
18. Nasim F., Masood S., Jaffar A., Ahmad U., Rashid M. Intelligent Sound-Based Early Fault Detection System for Vehicles // *Computer Systems Science and Engineering*. 2023. Vol. 46, no. 3. P. 3175–3190. DOI: 10.32604/csse.2023.034550.
19. Kawaguchi Y., Imoto K., Koizumi Y., Harada N., Niizumi D., Dohi K., Tanabe R., Purohit H., Endo T. Description and Discussion on DCASE 2021 Challenge Task 2: Unsupervised Anomalous Sound Detection for Machine Condition Monitoring under Domain Shifted Conditions // *Detection and Classification of Acoustic Scenes and Events 2021 : Proceedings*. Online, 15–19 Nov 2021. P. 186–190. ISBN 978-84-09-36072-7.
20. Ignjatovska A., Shishkovski D., Pecioski D. Classification of present faults in rotating machinery based on time and frequency domain feature extraction. *Vibroengineering Procedia*. 2023. Vol. 51. P. 22–28. DOI: 10.21595/vp.2023.23667.
21. Tuleski B. L., Yamaguchi C. K., Stefenon S. F., Coelho L. S., Mariani V. C. Audio-Based Engine Fault Diagnosis with Wavelet, Markov Blanket, ROCKET, and Optimized Machine Learning Classifiers // *Sensors*. 2024. Vol. 24. Art. 7316. DOI: 10.3390/s24227316.
22. Lüttenberg H., Bartelheimer C., Beverungen D. Designing Predictive Maintenance for Agricultural Machines // *Proceedings of the Twenty-Sixth European Conference on Information Systems (ECIS 2018)*. Portsmouth, UK, 2018. Research Paper No. 153. URL: https://aisel.aisnet.org/ecis2018_rp/153
23. Case Western Reserve University Bearing Data Center: веб-сайт. URL: <https://engineering.case.edu/bearingdatacenter> (дата звернення: 01.03.2026).
24. Cathy Siyu. Mechanical-datasets. [Dataset]. GitHub. URL: <https://github.com/cathysiyu/Mechanical-datasets> (дата звернення: 01.03.2026).

References

1. Aulin, V. V., Hrynkiv, A. V., Holovaty, A. O., Lysenko, S. V., Holub, D. V., Kuzyk, O. V., & Tykhyi, A. A. (2020). Methodological foundations of design and operation of intelligent transportation and manufacturing systems. Lysenko V.F. [in Ukrainian].
2. Matviienko, O. O., Aulin V.V. (2025). Classification of signal types and machining methods for intelligent assessment of the technical mill of mobile machines for agro-industrial production. *Central Ukrainian Scientific Bulletin. Technical Sciences*, (11(42)_II), 298-312. [https://doi.org/10.32515/2664-262X.2025.11\(42\).2.298-312](https://doi.org/10.32515/2664-262X.2025.11(42).2.298-312) [in Ukrainian].
3. Matviienko, O. O., Aulin V.V. (2025). Status and development directions of data architecture for intelligent assessment of the technical condition of mobile machines of agro-industrial enterprises. *Central Ukrainian Scientific Bulletin. Technical Sciences*, (12(43)_I), 227-237. [https://doi.org/10.32515/2664-262X.2025.12\(43\).1.227-237](https://doi.org/10.32515/2664-262X.2025.12(43).1.227-237) [in Ukrainian].
4. Loutas, T. H., Roulias, D., Pauly, E., & Kostopoulos, V. (2011). The combined use of vibration, acoustic emission and oil debris on-line monitoring towards a more effective condition monitoring of rotating machinery. *Mechanical Systems and Signal Processing*, 25(4), 1339–1352. <https://doi.org/10.1016/j.ymssp.2010.11.007>
5. Oyedaja, K. O. (2012). Diagnostics of bearing defects using vibration signal. *International Journal of Computer and Electrical Engineering*, 4(6), 821–825. <https://doi.org/10.7763/IJCEE.2012.V4.612>
6. Sen, A., Majumder, M. C., Mukhopadhyay, S., & Biswas, R. K. (2017). Condition monitoring of rotating equipment considering the cause and effects of vibration: A brief review. *International Journal of Modern Engineering Research (IJMER)*, 7(1), 36–49.
7. Ciaburro, G., & Iannace, G. (2022). *Machine-learning-based methods for acoustic emission testing: A review*. *Applied Sciences*, 12(20), 10476. <https://doi.org/10.3390/app122010476>
8. Da Silva, R. R., Da S. Costa, E., De Oliveira, R. C. L., & Mesquita, A. L. A. (2017). Fault diagnosis in rotating machine using full spectrum of vibration and fuzzy logic. *Journal of Engineering Science and Technology*, 12(11), 2952–2964.
9. Saberi, A. N., Belahcen, A., Sobra, J., & Vaimann, T. (2022). LightGBM-based fault diagnosis of rotating machinery under changing working conditions using modified recursive feature elimination. *IEEE Access*, 10, 81910–81925. <https://doi.org/10.1109/ACCESS.2022.3195939>
10. Zhang, B., Li, W., Tong, Z., & Zhang, M. (2017). *Bearing fault diagnosis under varying working condition based on domain adaptation* (arXiv:1707.09890). arXiv.
11. Wang, Q., Michau, G., & Fink, O. (2019). Domain adaptive transfer learning for fault diagnosis. In *2019 Prognostics and System Health Management Conference (PHM-Paris)*. IEEE. <https://doi.org/10.1109/PHM-Paris.2019.00054>
12. Zhang, X., & Gu, G. (2024). Fault diagnosis for rolling bearings under complex working conditions based on domain-conditioned adaptation. *Machines*, 12(11), 787. <https://doi.org/10.3390/machines12110787>

13. Zhang, Q., Lv, Z., Hao, C., Yan, H., & Fan, Q. (2024). Intelligent fault diagnosis of bearings in unsupervised dynamic domain adaptation networks under variable conditions. *IEEE Access*, 12, 82911–82925. <https://doi.org/10.1109/ACCESS.2024.3413087>
14. Zhong, X. (2024). Failure mechanism information-assisted multi-domain adversarial transfer fault diagnosis model for rolling bearings under variable operating conditions. *Sensors*, 24(3), 1036. <https://doi.org/10.3390/s24031036>
15. Li, X., Wang, J., Wang, J., Wang, J., Li, Q., Yu, X., & Chen, J. (2025). Research on unsupervised domain adaptive bearing fault diagnosis method based on migration learning using MSACNN-IJMMD-DANN. *Machines*, 13, 618. <https://doi.org/10.3390/machines13070618>
16. Jalayer, M., Kaboli, A., Orsenigo, C., & Vercellis, C. (2022). Fault detection and diagnosis with imbalanced and noisy data: A hybrid framework for rotating machinery. *Machines*, 10, 237. <https://doi.org/10.3390/machines10040237>
17. Duman, T. B., Bayram, B., & İnce, G. (2020). Acoustic anomaly detection using convolutional autoencoders in industrial processes. In *Advances in Intelligent Systems and Computing* (Vol. 1028, pp. 397–406). Springer. https://doi.org/10.1007/978-3-030-20055-8_41
18. Nasim, F., Masood, S., Jaffar, A., Ahmad, U., & Rashid, M. (2023). Intelligent sound-based early fault detection system for vehicles. *Computer Systems Science and Engineering*, 46(3), 3175–3190. <https://doi.org/10.32604/csse.2023.034550>
19. Kawaguchi, Y., Imoto, K., Koizumi, Y., Harada, N., Niizumi, D., Dohi, K., Tanabe, R., Purohit, H., & Endo, T. (2021). Description and discussion on DCASE 2021 Challenge Task 2: Unsupervised anomalous sound detection for machine condition monitoring under domain shifted conditions. In *Proceedings of the Detection and Classification of Acoustic Scenes and Events 2021 (DCASE 2021 Workshop)* (pp. 186–190).
20. Ignjatovska, A., Shishkovski, D., & Pecioski, D. (2023, October 20–21). *Classification of present faults in rotating machinery based on time and frequency domain feature extraction*. *Vibroengineering Procedia*, 51, 22–28. <https://doi.org/10.21595/vp.2023.23667>
21. Tuleski, B. L., Yamaguchi, C. K., Stefenon, S. F., Coelho, L. S., & Mariani, V. C. (2024). Audio-based engine fault diagnosis with wavelet, Markov blanket, ROCKET, and optimized machine learning classifiers. *Sensors*, 24, 7316. <https://doi.org/10.3390/s24227316>
22. Lüttenberg, H., Bartelheimer, C., & Beverungen, D. (2018). Designing predictive maintenance for agricultural machines. In *Proceedings of the Twenty-Sixth European Conference on Information Systems (ECIS 2018)*. https://aisel.aisnet.org/ecis2018_rp/153
23. Case Western Reserve University. (n.d.). *Bearing Data Center*. Case School of Engineering. <https://engineering.case.edu/bearingdatacenter>
24. Cathy Siyu. (n.d.). *Mechanical-datasets* [Dataset]. GitHub. <https://github.com/cathysiyu/Mechanical-datasets>

Oleksandr Matvienko, Assoc. Prof., PhD tech. sci., **Oleksander Livitskyi**, PhD tech. sci., **Andriy Gryniv**, PhD tech. sci.

Central Ukrainian National Technical University, Kropyvnytskyi, Ukraine

Methodology for Selecting Robust Vibration Diagnostics Models Under Domain Shift and Label Noise

This paper addresses a practical issue in vibration-based condition monitoring of mobile machines: diagnostic models that perform well on training data often degrade in the field due to domain shift (changes in operating regime, load, sensor mounting, and test-rig vs. real operation) and label noise (imprecise service records, ambiguous transition states, and human annotation errors).

We propose a conditional model-selection methodology formulated as a decision matrix that links domain-shift type and severity, labeled data volume and expected label reliability, and deployment constraints (validation feasibility and reproducibility requirements) to an appropriate class of models and training strategies.

The methodology is validated on two public scenarios with injected symmetric label noise: a moderate severity shift on the CWRU bearing dataset and a severe cross-condition shift on the SEU gearbox dataset. For the moderate shift, a two-stage robustification with confidence-based pseudo-label filtering achieves the best performance under high label noise, reaching Macro-F1 = 0.9368 at 20% noise. For the severe shift, a classical pipeline with simple feature-level domain alignment remains the most stable option, while ensemble strategies provide lower and less consistent gains. The results yield actionable guidance for integrating robust diagnostics into an intelligent technical service system under real-world constraints.

mobile machines, vibration diagnostics, domain shift, label noise, robust learning, pseudo-labeling, domain adaptation, rolling bearings; gearbox

Одержано (Received) 26.02.2026

Прорецензовано (Reviewed) 02.03.2026

Прийнято до друку (Approved) 12.03.2026