

Національний лісотехнічний університет України
(повне найменування вищого навчального закладу)

Навчально-науковий інститут комп'ютерних наук
та інформаційних технологій
(повне найменування інституту, назва факультету (відділення))

Кафедра комп'ютерних наук
(повна назва кафедри (предметної, циклової комісії))

Магістерська кваліфікаційна робота

другий (магістерський)

(рівень вищої освіти)

на тему: «Застосування методів машинного навчання для розроблення системи попереджень про втрати лісового покриву Львівщини на основі супутникових знімків»

Виконав студент 6 курсу, групи КН-63м

спеціальності: 122 “Комп'ютерні науки”
(шифр і назва напрямку підготовки спеціальності)

Пастернак С. Я.
(прізвище, ініціали)

Керівник:

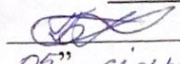
доц. Часковський О.Г.
(прізвище, ініціали)

Рецензент: Карашецький В. П.
(прізвище, ініціали)

Львів-2024

НАЦІОНАЛЬНИЙ ЛІСОТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
(повне найменування вищого навчального закладу)

ННІ комп'ютерних наук та інформаційних технологій
Кафедра Комп'ютерних наук
Рівень вищої освіти другий (магістерський)
Спеціальність 122 "Комп'ютерні науки"

ЗАТВЕРДЖУЮ:
Завідувач кафедри
комп'ютерних наук
 Борецька І.Б.
„05” січня 2024 р.

**ЗАВДАННЯ
НА ДИПЛОМНУ РОБОТУ СТУДЕНТУ**

Пастернаку Степану Ярославичу
(прізвище, ім'я, по батькові)

1. Тема магістерської роботи: Застосування методів машинного навчання для розроблення системи попереджень про втрати лісового покриву Львівщини на основі супутникових знімків

керівник роботи: доц. Часковський Олег Григорович, к.с.-г.н, доцент,
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом університету від “13” лютого 2023 року № С-49

2. Термін подання студентом проекту (роботи) 5 січня 2024 р.

3. Вихідні дані до проекту (роботи). Документація «Google Earth Engine», Java Script та «Python», матеріали мережі Інтернет, що стосуються, роботи.

4. Зміст пояснювальної записки (перелік питань, які потрібно розробити)

Розділ 1 Стан проблемної області

Розділ 2. Інформаційне забезпечення

Розділ 3. Математичне забезпечення

Розділ 4. Програмне забезпечення

Розділ 5. Розроблення стартап проекту

Висновки

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)

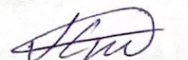
Підготовка матеріалу до доповіді (схеми дерева рішень для класифікації супутникових знімків, схема побудови системи попередження про втрати лісового покриву, приклади карт наземного покриву та втрат лісового покриву, знімки з екрана створення та передачі повідомлень).

6. Дата видачі завдання “15” лютого 2023 року

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Етапи магістерської роботи	Термін виконання етапів роботи	Примітка
1	Огляд літературних та інших джерел згідно досліджуваної теми.	14.02.23 – 27.03.23	виконано
2	Аналіз досліджуваної теми та вибір відповідних варіантів її розробки.	4.04.23 – 24.04.23	виконано
3	Постановка задачі та її формалізація.	25.04.23 – 1.05.23	виконано
4	Вибір та обґрунтування методів і засобів проведення дослідження.	2.05.23 – 22.05.23	виконано
5	Розроблення концептуальної схеми реалізації завдання.	23.05.23 – 29.05.23	виконано
6	Програмна реалізація завдання.	30.05.23 – 30.08.23	виконано
7	Тестування програмного продукту та отриманих результатів	31.18.23 – 30.10.23	виконано
8	Розробка пояснювальної записки та презентації дипломної роботи	31.10.23 – 9.01.24	виконано

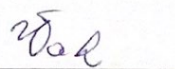
Студент КН-63м


(підпис)

Пастернак С.Я.

(прізвище та ініціали)

Керівник роботи


(підпис)

доц. Часковський О.Г.

(прізвище та ініціали)

РЕФЕРАТ

магістерської роботи студентки групи КН-63м Пастернака Степана Ярославовича на тему: **Застосування методів машинного навчання для розроблення системи попереджень про втрати лісового покриву Львівщини на основі супутникових знімків**. Обсяг роботи: 51 стор.; кількість ілюстрацій: 10; додатків: 1; використаних джерел: 8.

Магістерська робота присвячена розробці актуальної теми сьогодення, а саме розробка програмного забезпечення для виявлення втрат лісового покриву та швидкого оповіщення відповідних органів. Для реалізації використано програми Google Earth Engine та мову програмування «Java Script» для побудови системи спостереження. Для опрацювання супутникових знімків з метою виявлення втрат лісового покриву використано методи машинного навчання та їх використання в середовищі Google Earth Engine. Окрім того, розроблена система для оповіщення про місця втрат лісового покриву даних зйомки. Обґрунтовано актуальність вибраного проєкту. Система розроблялася із врахуванням потреб зацікавлених сторін, установ та громадськості.

Ключові слова: Google Earth Engine, програмне забезпечення, суцільні рубки, Java Script, алерт-систем, Python.

ABSTRACT

The master's thesis is devoted to the development of a current topic, namely the development of software for detecting the loss of forest cover and quickly notifying the relevant authorities. Google Earth Engine programs and Java Script programming language were used to build the surveillance system. Machine learning methods and their use in the Google Earth Engine environment were used to process satellite images in order to detect the loss of forest cover. In addition, a system has been developed for notifying the locations of loss of forest cover from survey data. The relevance of the selected project is substantiated. The system was developed taking into account the needs of stakeholders, institutions and the public.

Keywords: Google Earth Engine, software, clear cutting, Java Script, alert system, Python.

ТЕХНІЧНЕ ЗАВДАННЯ

Розробити систему попередження втрат лісового покриву на основі супутникових знімків для цього:.

Підібрати супутникові знімки достатньої просторової та спектральної роздільної здатності;

- Визначити метод машинного навчання, придатний для інтерпретації втрат лісового покриву за супутниковими знімками;
- Підібрати тестові ділянки для навчання алгоритму машинного навчання;
- Провести класифікацію знімків з метою виявлення втрат лісового покриву;
- Визначити точність карт втрат лісового покриву;
- Організувати інформування через електронну пошту зацікавлених осіб про місця втрат лісового покриву.
- Надати рекомендації із використання знімків для функціонування системи створення карти втрат лісового покриву та порівняння її з дозвільними документами.

ЗМІСТ

<u>РЕФЕРАТ</u>	1
<u>ТЕХНІЧНЕ ЗАВДАННЯ</u>	2
<u>Вступ</u>	6
<u>РОЗДІЛ 1. СТАН ПРОБЛЕМНОЇ ОБЛАСТІ</u>	8
<i><u>1.1. Актуальність створення систем попередження про втрати деревного покриву</u></i>	8
<i><u>1.2. Технології машинного НАВЧАННЯ та їх роль для опрацювання супутникових знімків</u></i>	9
<u>РОЗДІЛ 2. ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ</u>	11
<i><u>2.1 Інформаційне забезпечення алерт-системи</u></i>	11
<i><u>2.1. Коротка характеристика мультиспектральної знімальної системи LANDSAT</u></i>	11
<i><u>2.2. Коротка характеристика радіолокаційної знімальної системи ERS</u></i>	15
<u>РОЗДІЛ 3. МАТЕМАТИЧНЕ ЗАБЕЗПЕЧЕННЯ</u>	18
<i><u>3.1. Методи машинного навчання для класифікації супутникових знімків</u></i>	18
<i><u>3.2. Показники часових рядів матеріалів дистанційних спостережень</u></i>	28
<i><u>3.4. Компоненти часових рядів матеріалів дистанційних методів спостережень.</u></i>	30
<i><u>3.3. Близький до реального часу моніторинг</u></i>	30
<u>РОЗДІЛ 4. ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ</u>	34
<i><u>4.1 Застосування і основні принципи роботи Google Earth Engine</u></i>	34
<i><u>4.2 Python і Java Script та їх застосування в Google Earth Engine</u></i>	35
<i><u>4.3. Google monitoring для генерування повідомлень</u></i>	40
<i><u>4.4. Особливості роботи системи попереджень про втрати лісового покриву на території Львівщини.</u></i>	47
<u>РОЗДІЛ 5. РОЗРОБЛЕННЯ СТАРТАП-ПРОЕКТУ</u>	49
<u>5.1 Опис ідеї проекту</u>	49
<i><u>5.2 Розроблення ринкової стратегії</u></i>	50
<i><u>5.3. Вимоги до технічного та програмного забезпечення</u></i>	53
<u>ВИСНОВКИ</u>	55
<u>СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ</u>	57
<u>ДОДАТКИ</u>	58
<i><u>Додаток А</u></i>	58

Вступ

Супутникові дані дають можливість покрити всі ділянки України та виявляти будь-які зміни в режимі реального часу. Моделі штучного інтелекту дають можливість уточнити, скільки стовбурів дерев зникло, і для цього не треба нікуди виходити з офісу, навіть не треба нічого рахувати, бо ці моделі самі прозондують та порахують кожен стовбур, який зник. Цей метод поєднує супутникові дані, державні дані та штучний інтелект для розв'язання задач моніторингу лісів. Алгоритм дій наступний – спочатку обробка супутникових знімків, далі – формування карт зникнення лісів, порівняння їх із лісорубними квитками і виявлення сумнівних рубок, котрі за бажання і можливості можна перевірити наочно.

З моменту запуску першого супутника наземного спостереження Landsat-1 у 1972 році для класифікації пікселів на зображеннях Thematic Mapper (TM) використовувалося багато алгоритмів машинного навчання. Методи класифікації варіюються від параметричних контрольованих алгоритмів класифікації, таких як максимальна правдоподібність, неконтрольованих алгоритмів, таких як ISODAT і-кластеризація k-середніх, до алгоритмів машинного навчання, таких як штучні нейронні, дерева рішень, опорні векторні машини та класифікатори ансамблів. В останні роки були запропоновані різні алгоритми класифікації ансамблю. Найбільш широко використовуваним алгоритмом класифікації ансамблю є випадковий ліс. Класифікатор Random Forest використовує початкове агрегування для формування ансамблю класифікаційного та індукційного дерева, подібного до деревних класифікаторів.

Актуальність роботи полягає у вирішенні проблеми об'єктивної та оперативної оцінки лісового покриття, його втрат та виявлення їх походження, а при невідповідності з дозвільними документами і швидкого реагування на них.

Предметом досліджень є методи машинного навчання, придатні для якомога точнішого дешифрування втрат деревного покриття за супутниковими знімками.

Об'єкт дослідження – лісові насадження та їх динаміка під впливом природних та антропогенних факторів;

Мета досліджень побудувати систему розпізнавання втрат лісового покриття за допомогою методів машинного навчання для інтерпретації супутникових знімків.

Для цього необхідно виконати наступні **завдання**:

- Підібрати супутникові знімки достатньої просторової та спектральної роздільної здатності;
 - Визначити метод машинного навчання, придатний для інтерпретації втрат лісового покриву за супутниковими знімками;
 - Підібрати тестові ділянки для навчання алгоритму машинного навчання;
 - Провести класифікацію знімків з метою виявлення втрат лісового покриву;
 - Побудувати систему оповіщення про втрати лісового покриву;
 - Надати рекомендації із використання знімків для функціонування системи створення карти втрат лісового покриву та порівняння її з дозвільними документами.

Наукова новизна роботи полягає у тому, що вперше для території досліджень використані супутникові знімки та алгоритми машинного навчання для визначення втрат лісового покриву. Також передбачається створення системи оповіщення про втрати лісового покриву через електронну пошту зацікавлених сторін.

Практична значимість роботи полягає у можливості використання такої системи для контролюючих та урядових органів, а також для широкого кола громадськості.

РОЗДІЛ 1. СТАН ПРОБЛЕМНОЇ ОБЛАСТІ

1. Актуальність створення систем попередження про втрати деревного покриву

Створення систем попередження про втрати деревного покриву має надзвичайну актуальність у сучасному світі з огляду на кілька важливих причин:

Зміна клімату: Ця проблема вже має серйозний вплив на екосистеми та людей по всьому світу. Втрати деревного покриву призводять до зменшення здатності лісів поглинати вуглець, що призводить до збільшення викидів парникових газів та глобального потепління.

Втрата біорізноманітності: Ліси є домівками для безлічі видів рослин і тварин, і їх втрата призводить до зменшення біорізноманітності. Це може вплинути на продовольчу безпеку та природний баланс екосистем.

Знищення природних ресурсів: Ліси забезпечують важливі природні ресурси, такі як деревина, лікарські рослини, інгредієнти харчування та інші матеріали. Втрата деревного покриву може призвести до економічних проблем та знищення природних екосистем.

Повені та зсуви ґрунту: Ліси функціонують як природні об'єкти водосховища та стабілізації ґрунту. Втрата лісів може призвести до частіших повеней та зсувів ґрунту, що загрожує життям та майну населення.

Соціальні наслідки: Втрата деревного покриву може призвести до втрати робочих місць для місцевого населення, зокрема для тих, хто займається лісозаготівлею та сільським господарством.

Глобальний інтерес: Охорона лісів є глобальним питанням, і багато країн та міжнародні організації приділяють значну увагу проблемам лісозаготівлі та втрати деревного покриву. Створення систем попередження дозволяє вчасно виявляти та втручатися в процеси втрати лісів.

Технологічні можливості: Завдяки сучасним технологіям, таким як супутникове зображення та машинне навчання, стало можливим ефективно відстежувати та передбачати втрати деревного покриву з високою точністю, що

робить створення систем попередження більш реалістичним та ефективним завданням.

Загалом, створення систем попередження про втрати деревного покриву є надзвичайно актуальним завданням, оскільки це дозволяє зберігати природні ресурси, зменшувати вплив зміни клімату, підтримувати біорізноманітність та забезпечувати сталість екосистем для майбутніх поколінь.

1.2. Технології машинного навчання та їх роль для опрацювання супутникових знімків

Технологія машинного навчання є підгалуззю штучного інтелекту (AI), яка спрямована на розробку алгоритмів та моделей, що дозволяють комп'ютерам навчатися на основі даних та приймати прогнози або рішення на основі даних. Вона включає в себе використання статистичних технік, які дозволяють комп'ютерам автоматично покращувати свою продуктивність у виконанні певної задачі завдяки досвіду, без явного програмування. Машинне навчання знайшло застосування в різних галузях, включаючи розпізнавання зображень, обробку природної мови, медичну діагностику та інше.

У контексті створення системи оповіщення про руйнування лісів на основі супутникових зображень технологія машинного навчання грає важливу роль:

1. Збір даних: Перший крок у створенні такої системи включає в себе збір великого набору даних супутникових зображень, які містять як руйнування лісів, так і неруйновані території. Зазвичай ці зображення отримуються від супутників, які обертаються навколо Землі, і можуть надавати цінну інформацію про стан лісів.

2. Передобробка даних: Супутникові зображення часто містять шум та різні артефакти. Алгоритми машинного навчання використовуються для передпоцесингу та очищення даних, включаючи видалення хмар, корекцію атмосферних ефектів і покращення якості зображення.

3. Видобуток ознак: Моделі машинного навчання повинні видобувати відповідні ознаки або патерни з супутникових зображень. Ознаки можуть включати

щільність рослинності, типи земельного покриття та зміни в лісовому покриві з часом. Зазвичай для цієї мети використовують згорткові нейронні мережі (CNN), оскільки вони добре підходять для завдань аналізу зображень.

4. Маркування та навчання: Набір даних потрібно маркувати, де кожне зображення класифікується як руйнування лісів чи неруйнування лісів. Моделі машинного навчання, такі як алгоритми навчання з учителем, навчаються на цьому маркованому наборі даних для вивчення патернів та відносин між ознаками та руйнуванням лісу.

5. Розробка моделі: Дослідники і фахівці з обробки даних розробляють моделі машинного навчання, такі як випадкові ліси, метод опорних векторів чи глибокі моделі навчання, як CNN, для прогнозування руйнування лісу на основі видобутих ознак. Ці моделі навчаються за допомогою маркованого набору даних і можуть робити прогнози для нових, немаркованих супутникових зображень.

6. Валідація та тестування: Результативність моделей машинного навчання оцінюється за допомогою окремого набору валідаційних і тестових даних, щоб переконатися, що вони добре узагальнюються до нових, невидимих зображень. Метрики, такі як точність, точність, відзив і F1-показник, часто використовуються для оцінки продуктивності моделі.

7. Генерація оповіщень: Як тільки розроблено модель машинного навчання, яка добре виконує своє завдання, її можна впровадити як частину системи оповіщення. Ця система безперервно обробляє надходячі супутникові зображення та позначає області з можливим руйнуванням лісу. Оповіщення можуть бути відправлені відповідним органам, організаціям з охорони природи або іншим учасникам у реальному часі або на регулярній основі.

8. Постійне покращення: Моделі машинного навчання можна подальше покращувати, перетренувавши їх на нових даних та адаптувавши їх до зміни умов. Це забезпечує, що система оповіщення залишається точною та ефективною з плином часу.

РОЗДІЛ 2. ІНФОРМАЦІЙНЕ ЗАБЕЗПЕЧЕННЯ

2.1 Інформаційне забезпечення алерт-системи

Сучасна інформація про стан лісових покривів здобувається за допомогою технік дистанційного зондування. Ці техніки поділяються на активні та пасивні методи. Пасивні методи включають фотографічне, телевізійне, спектрометричне, та сканерне зондування, які використовують природне або відбите випромінювання Землі, освітлене сонячним світлом. Ці методи дозволяють досягти високої контрастності зображень та чіткіше виділяти об'єкти. Активні методи, як-от радарне знімання, самостійно освітлюють цільову зону високочастотними імпульсами, згодом реєструючи відбитий сигнал.

Радари з бічним оглядом, працюючи в мікрохвильовому діапазоні, революціонізували зйомку, дозволяючи проводити її в будь-який час і при будь-яких погодних умовах, оскільки вони проникають через хмари, дощ і туман. Такі системи особливо корисні в постійно хмарних регіонах, де традиційне аерофотознімання утруднене.

Використання мультиспектральних сканерів та радарів бічного огляду дозволяє ефективно здійснювати моніторинг та тематичне картографування Землі, але не може повністю замінити аерофотознімки через їхню доступність і простоту. Дистанційне зондування використовує видимі та невидимі діапазони спектру, вимагаючи технічної досконалості обладнання і глибоких знань від спеціалістів. Класифікація зображень може проводитися візуально або автоматизовано, кожен з методів має свої переваги та обмеження.

2.1. Коротка характеристика мультиспектральної знімальної системи LANDSAT

Наше дослідження ґрунтується на зображеннях, здобутих завдяки серії штучних супутників LANDSAT. Ці сателіти перебувають на сонячно синхронних полярних орбітах, які охоплюють від 81° північної до 81° південної широти. Така конфігурація

орбіти дозволяє регулярно знімати одну й ту саму територію в однаковий час. Кожен супутник забезпечує покриття земної поверхні шириною 183 км. Повний оборот зйомки земної поверхні здійснюється кожні 16 днів, причому Landsat 7 слідує за Landsat 5 із затримкою у 8 днів. Детальні технічні характеристики супутників Landsat представлені в таблиці 2.1.

Таблиця 2.1. – Технічні характеристики супутників Landsat

Супутник	Дата запуску	Закінчення дії	Висота зйомки, км	Кут нахилу, град.	Час польоту навколо земної кулі, днів	Тип сенсора
Landsat-1	23.07.1972	06.01.1978	920	99,2	18	RBV, MSS
Landsat-2	22.06.1975	05.02.1982	920	99,2	18	RBV, MSS
Landsat-3	05.03.1978	31.03.1983	920	99,2	18	RBV, MSS
Landsat-4	16.07.1982	07.1987	705	98,2	16	MSS, TM
Landsat-5	01.03.1985	–	705	98,2	16	MSS, TM
Landsat-6	невдалий запуск					
Landsat-7	15.04.1999	–	705	98,2	16	ETM+

Супутники серії Landsat від 1 до 3 були оснащені камерами типу RBV (Return Beam Vidicon) та мультиспектральними сканерами MSS (Multispectral Scanner). Знімання, яке проводили Landsat 1 та 2, відбувалося у видимому та ближньому інфрачервоному діапазонах спектру. Починаючи з третього покоління, супутники також здатні фіксувати зображення в інфрачервоному тепловому діапазоні. З 1982 року, нові супутники Landsat мають додатковий модуль TM (Thematic Mapper), який працює разом із MSS. Landsat 7 використовує поліпшену версію цього сенсора, ETM+ (Enhanced Thematic Mapper). Відмінність між TM та ETM+ полягає в тому, що ETM+ включає додатковий панхроматичний канал із роздільною здатністю 15 метрів та 6-й

канал (термальний), який поділений на 2 частини (див. табл. 2.2).

Таблиця 2.2. – Типи камер Landsat та спектральні діапазони зйомки

Канали	Тип камери MSS		Тип камери TM&ETM		Тип камери ETM+	
	спектральний діапазон, мкм	розмір пікселя, м	спектральний діапазон, мкм	розмір пікселя, м	спектральний діапазон, мкм	розмір пікселя, м
1	–	–	0,45–0,52	30	–	–
2	–	–	0,52–0,60	30	–	–
3	–	–	0,63–0,69	30	–	–
4	0,5–0,6	80	0,76–0,90	30	–	–
5	0,6–0,7	80	1,55–1,75	30	–	–
6 (терм.)	0,7–0,8	80	10,42–12,50	120(ETM 60)	–	–
7	0,8–1,1	80	2,08–2,35	30	–	–
8 (pan)	–	–	–	–	0,52–0,9	15

Як вказано у таблиці 2.2, камери кожного типу включають фільтри, що пропускають світло у певних спектральних діапазонах, дозволяючи отримати інформацію з конкретними спектральними характеристиками. Це значно спрощує дешифрування космічних зображень та збільшує кількість отриманої інформації.

Підготовка до вибору знімків для дешифрування вимагає усвідомленого підходу до завдань дослідження та методів дешифрування. Одним з ключових критеріїв є

вибір просторової розрізняювальної здатності знімків, яка найкраще підходить для цілей дослідження. При виборі матеріалів для масштабних досліджень також враховують економічну доцільність та можливість покриття всієї території, а також доступність архівних зображень.

Знімки Landsat різних поколінь, доступні з 1975 до 2013 року, вважаються одними з найкращих для таких досліджень, оскільки вони надають мультиспектральні зображення від видимого до середнього інфрачервоного діапазону, що є ідеальним для вивчення рослинного покриття, та панхроматичні зображення з вдвічі вищою просторовою розрізняювальною здатністю.

Проте, при виборі знімків слід звертати увагу на відсутність хмарності, оскільки для наукових досліджень придатні тільки знімки з 30% і меншим покриттям хмарами. Також слід враховувати можливе зміщення сканів на знімках Landsat-8, що може вимагати додаткової роботи з геометричним коригуванням.

Після ретельної оцінки архівних знімків Landsat на предмет хмарності та геометричних спотворень, формується база даних для дистанційного зондування Землі (табл. 2.3). Зібрані матеріали потім об'єднуються у мультиспектральне зображення, яке перетворюється в багатозональне кольорове зображення для подальшого дешифрування. Кінцеве зображення буде в кольоровому форматі з просторовою розрізняювальною здатністю 28,5 м.

Таблиця 2.3. – Банк даних матеріалів дистанційного зондування Землі для дослідження змін землекористування

№ п/п	Сцена	Супутник	Рік зйомки
12	I T51250252010152M0P00 for	I 5	2010
10	I T51250251000122VYV02 for	I 5	1000
20	I T51250251020722K1C00 for	I 5	1020
21	I T51250251026216VYV01 for	I 5	1026
22	I T51250251026221VYV05 for	I 5	1026
23	I T51250251026230VYV05 for	I 5	1026
24	I T51250251026121VYV02 for	I 5	1026
25	I T41260251022200VYV02 for	I 4	1022
26	I T41260251022221VYV05 for	I 4	1022
27	I E71260252006252A C000 for	I 7	2006
28	I E71260252002266E0C01	I 7	2002
29	I E71260252002106A C000	I 7	2002
30	I E71260252002071S0C00 for	I 7	2002
31	I E71260252000050E0C00 for	I 7	2000

20	I E71850252005261A SN00 for	I 7	2005
22	I E71850252003147A SN00 for	I 7	2002
24	I E7185025 02520000502 for	I 7	2000
25	I C81860252013172I SN00 (1) for	I 8	2012
26	I C81860252013157I SN00 for	I 8	2012
27	I C81850252013150I SN00 (1) for	I 8	2012

2.2. Коротка характеристика радіолокаційної знімальної системи ERS

Радіолокаційні знімальні системи, як активні системи, формують зображення за допомогою імпульсів, відбитих від поверхні, яку опромінювала сама система. Це принципово відрізняє їх від пасивних систем, які фіксують лише відбиті чи випромінені поверхнею природні сигнали. Радіолокаційні системи сканують один бік від носія, зазвичай перпендикулярно до лінії польоту. Це особливо важливо для систем бічного огляду з реальною або синтезованою апертурою антени.

Знімальна система ERS-2 SAR використовує технологію синтезованої апертури для забезпечення високої роздільної здатності, знімаючи земну поверхню з висоти 782 км. Завдяки цьому супутник забезпечує широкий огляд смуги на поверхні Землі шириною 100 км та довжиною хвиль у С-діапазоні, що дозволяє отримувати зображення з високою просторовою розрізнявальною здатністю.

Однак, радіолокаційні знімки, незалежно від погодних умов, часу доби чи хмарності, потребують комплексної попередньої обробки перед використанням. Це включає радіометричну, геометричну та сенсорну корекцію, щоб забезпечити точне відображення поверхні. Також часто використовується speckle-filtering для зменшення шуму та поліпшення якості зображення.

З метою статистичної оцінки ефективності класифікації, для RandomForest було розраховано показник загальної точності, Каппа-коефіцієнт та матрицю змішування, що дозволяє глибше аналізувати якість розподілу даних на класи. Точність розподілу виявилася високою, досягаючи до 98% у деяких випадках.

Загалом, радіолокаційні системи надають унікальну здатність здобування даних про земну поверхню, що важко досяжна для традиційних методів дистанційного

зондування. Втім, для повноцінного використання цих даних потрібно провести ретельну попередню обробку та статистичну оцінку.

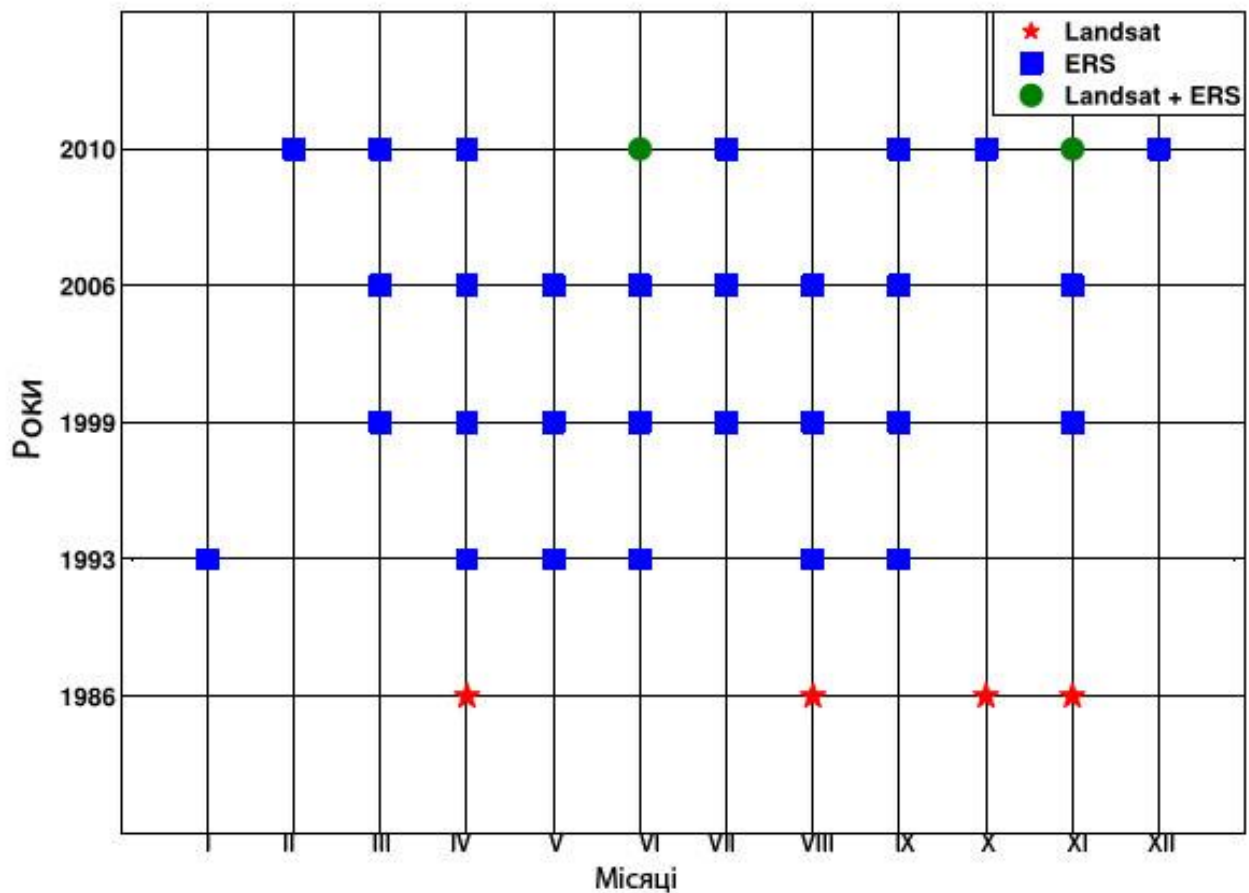


Рис. 2.1. Наявні оптичні та радарні супутникові знімки.

Отримані дані дистанційного зондування Землі є цінним джерелом інформації для різноманітних досліджень та застосувань. Однак, для того, щоб ці дані були корисними, їх необхідно інтерпретувати та перетворити в зрозумілі та доступні для аналізу формати. Цей процес часто включає використання спеціалізованого програмного забезпечення та розробку класових сигнатур на основі польових досліджень.

Інтерпретація даних в спеціалізованих програмних продуктах дозволяє обробляти, аналізувати та візуалізувати отримані зображення чи інші види даних. Це може включати в себе корекцію зображень, класифікацію, аналіз текстур та інші види обробки, які допомагають перетворити сирий набір даних у корисну інформацію.

Класові сигнатури, створені на основі польових досліджень, є важливим

елементом в процесі класифікації даних дистанційного зондування. Вони використовуються як еталони для ідентифікації та класифікації різних типів поверхонь на Землі, таких як водойми, ліси, сільськогосподарські угіддя, міські території тощо. Класові сигнатури визначають характерні особливості кожного типу поверхні, які потім використовуються для автоматичного розпізнавання цих типів на знімках дистанційного зондування.

Завдяки поєднанню спеціалізованого програмного забезпечення та точних класових сигнатур, можна досягти детальної та точної інтерпретації даних дистанційного зондування. Це відкриває широкі можливості для аналізу змін у ландшафтах, моніторингу навколишнього середовища, планування використання земель та багато інших застосувань.

РОЗДІЛ 3. МАТЕМАТИЧНЕ ЗАБЕЗПЕЧЕННЯ

3.1. Методи машинного навчання для класифікації супутникових знімків

Для тлумачення супутникових даних за спектральними особливостями розроблено методики аналізу та тлумачення знімків, які фокусуються на структурній інформації, тобто класифікація відбувається не за окремими пікселями, а за цілісними об'єктами. Для розпізнавання образів на знімках Landsat застосовано як попівсельну, так і об'єктно-орієнтовану класифікацію, тоді як для знімків ERS+Landsat використовувались ієрархічні об'єктно-орієнтовані методики.

Класифікація мультиспектральних зображень вже давно привертає увагу фахівців у галузі дистанційного зондування, адже результати класифікації становлять основу для різноманітних екологічних та соціально-економічних досліджень. Класифікація пікселів є ключовим етапом у аналізі зображень Thematic Mapper (TM). Для підвищення точності класифікації вчені впроваджують новітні підходи та методології. Втім, викликами залишаються численні чинники, такі як особливості досліджуваної місцевості, вибір дистанційних даних та методи обробки зображень, які можуть вплинути на успіх класифікації. Існує безліч методів для аналізу зображень Landsat TM, включаючи статистичні методи та техніки м'якого обчислення, такі як нейронні мережі, системи нечіткого висновку та нечіткі нейронні системи. Класичні статистичні методи, які використовуються для класифікації пікселів у мультиспектральних зображеннях, охоплюють класифікатор максимальної правдоподібності, класифікатор мінімальної відстані та різноманітні методи кластеризації. Класифікатор максимальної правдоподібності заснований на нормальних функціях розподілу для відбиття світла та обчислює середній вектор та коваріаційну матрицю для кожного класу за допомогою навчальних даних. Для призначення пікселя до певного класу використовуються апостеріорні ймовірності, обчислені за допомогою закону Байєса.

Використання методу RandomForest (RF) для дешифрування даних дистанційного зондування Землі виявилось ефективним завдяки його швидкості та високій точності класифікації, яка є стійкою до значних варіацій в даних. Незважаючи

на широке використання Random Forest у дослідженнях земельного покриття, його потенціал ще не повністю використаний спільнотою дистанційного зондування. У цій роботі проведено порівняння точності класифікації випадкового лісу з іншими поширеними алгоритмами, такими як максимальна правдоподібність, мінімальна відстань, дерево рішень, нейронні мережі та класифікатори машин опорних векторів.

Нейронні мережі, які є потужною та розумною альтернативою традиційним класифікаторам, забезпечують вищий рівень надійності та толерантності. Вони здатні безпосередньо проєціювати вхідний вектор спостереження на вихідну категорію, після завершення навчання. Нейронні мережі особливо ефективні для великих зображень і широко використовуються для класифікації пікселів у мультиспектральних зображеннях. Ряд досліджень, включаючи роботи Chen et al., Foody, Huang та Lippmann, Eberlein та ін., Cleeremans та ін., та Decatur, продемонстрували ефективність нейронних мереж у класифікації земельного покриття та аналізі даних. Кулкарні та Лулла розробили кілька моделей нейронних мереж, включаючи тришарові та чотиришарові моделі, які використовувалися як контрольовані класифікатори для розпізнавання пікселів за їхніми спектральними характеристиками на знімках Landsat. Алгоритми кластеризації, такі як split-merge, нечіткі K-means і методи на основі нейронних мереж, також використовувалися для мультиспектрального аналізу. Kulkarni та McCaslin застосували нейронні мережі для класифікації пікселів у мультиспектральних зображеннях та екстракції знань.

Машини опорних векторів (SVM) представляють собою контрольовану непараметричну статистичну техніку навчання, метою якої є визначення гіперплощини, що оптимально розділяє навчальні зразки на задану кількість класів. SVM працюють як бінарні класифікатори, призначаючи кожен тестовий зразок до одного з двох можливих класів. Для роботи з нелінійно роздільними даними, SVM використовують функції ядра для трансформації даних у вищий розмірний простір. В області дистанційного зондування SVM особливо цінуються за їхню здатність ефективно працювати з невеликими навчальними наборами, часто надаючи більш високу точність класифікації порівняно з традиційними методами.

Дерева рішень є ще одним видом алгоритмів класифікації, які, хоч і не користуються широкою популярністю у спільноті дистанційного зондування, все ж привабливі своєю здатністю обробляти ненормальні, неоднорідні та зашумлені дані. Класифікаційні дерева, запропоновані як альтернатива традиційним методам, використовувалися для класифікації мультиспектральних зображень, демонструючи ефективність у порівнянні з класифікатором максимальної правдоподібності.

Нещодавно, у класифікації земельного покриття використовуються ансамблеві методи, такі як випадковий ліс. Цей алгоритм базується на класифікаційних деревах і включає в себе багато класифікаційних дерев. Класифікація нового вектора ознак здійснюється за допомогою "голосування" кожного дерева у лісі. Випадковий ліс відзначається високою точністю, ефективністю на великих наборах даних та можливістю збереження структури для майбутнього використання. Метод випадкового лісу також важливий для класифікації даних з різних джерел, оскільки не лише непараметричний, але й дозволяє оцінювати значущість окремих змінних у класифікації.

Класифікатори дерева рішень відрізняються більшою ефективністю порівняно з одноетапними класифікаторами, приймаючи рішення на кількох рівнях, і відомі як багаторівневі класифікатори. Основні проблеми цих класифікаторів полягають у виборі найефективніших характеристик для поділу класів та умовах поділу для вузлів. Метою є побудова оптимального дерева для досягнення високої точності з мінімальними обчисленнями. Різні методи, включаючи ентропію та інші, використовуються для проектування дерев, з особливим акцентом на умовах зупинки, які можуть ґрунтуватися на порогових значеннях глибини або кількості випадків у вузлі.

ID3, одна з різновидів базового дерева рішень, використовує ентропію як умову розщеплення та загальну чистоту вузла як умову зупинки. Цей метод швидко вирощує точне дерево, використовуючи лише частину навчального набору для вирощування, а решту - для класифікації та подальшого вирощування. Такий підхід дозволяє досягти високої точності, але не гарантує

$$Gain(A) = Info(D) - Info_A(D) \text{ Де,}$$

$$Info_A(D) = \sum_{i=1}^n \frac{D_i}{D} \times Info(D)$$

Info(D) можна визначити, використовуючи одне з вищезгаданих рівнянь, що залежить від обраного методу поділу. InfoA(D) розраховується з використанням формули (3), де вага і-го поділу, а n - кількість різних значень атрибуту A. Атрибут A, що має найвищий приріст інформації, Gain(A), обирається для розбиття у вузлі N. Процес відбувається рекурсивно. Завдяки цьому підходу Квінлан зміг швидко створювати ефективні та точні дерева, не використовуючи великі навчальні набори, тим самим скорочуючи час і витрати на побудову. С4.5, що є нащадком Квінлана, - це алгоритм навчання під наглядом, який може використовувати як безперервні, так і дискретні атрибути, враховує набори даних із неповними даними та дозволяє призначати різні ваги для атрибутів, що може бути корисним для точнішого моделювання даних. Можна побудувати кілька дерев з С4.5 і використовувати їх в ансамблі для створення випадкового лісу.

Breiman запропонував концепцію пакетування, або «bootstrap aggregating». Ця ідея полягає у використанні кількох версій прогнозувача чи класифікатора, щоб прийняти остаточне рішення за допомогою більшості голосів серед цих прогнозувачів. Було виявлено, що зі збільшенням кількості прогнозувачів точність також зростає до певного порога, після чого вона знижується. Виявлення оптимальної кількості прогнозувачів може забезпечити найвищу точність. Пал та Матер оцінили ефективність класифікатора дерева рішень для класифікації земельного покриття, підвищивши точність класифікації дистанційних зображень за допомогою пакетування кількох дерев рішень. Випадкові ліси створюються за принципами пакетування та ID3. Кожне дерево у лісі вирощується, використовуючи випадкову підмножину навчального набору (з заміною), що відображає ідею ID3. Проте, кожен випадок у цьому наборі не використовується для вирощування дерева. Близько однієї третини вибірки не використовується для вирощування та вважається даними поза пакетом (ООВ). Додатково, не всі атрибути використовуються для побудови дерева. У кожному вузлі випадково оцінюється підмножина атрибутів. Дані ООВ

використовуються для оцінки частоти помилок класифікації, коли дерева додаються до лісу, і для оцінки важливості вхідних змінних. Після завершення лісу, класифікація може бути виконана за допомогою більшості голосів серед усіх дерев у лісі, що відображає ідею початкового агрегування.

Випадковий ліс також може оцінювати важливість змінних, використовуючи дані ООВ. Кожна змінна m випадково переставляється, і переставлені випадки ООВ знову класифікуються через дерево. Різниця між кількістю правильно класифікованих випадків із переставленими та непереставленими даними дає оцінку важливості змінної m . Ці оцінки різняться для кожного дерева, але середнє значення по всім деревам дає приблизну оцінку важливості кожної змінної. Random Forest був реалізований за допомогою пакета програмного забезпечення на мові R і використаний для аналізу зображень Landsat.

Основою методу RandomForest є класифікаційні та регресійні дерева (CART), розроблені Брайманом та іншими. Це бінарні дерева рішень. Найпростіший приклад CART представлений на рисунку 1.10. CART можуть використовуватися як для визначення лінійної регресії, так і для класифікації.

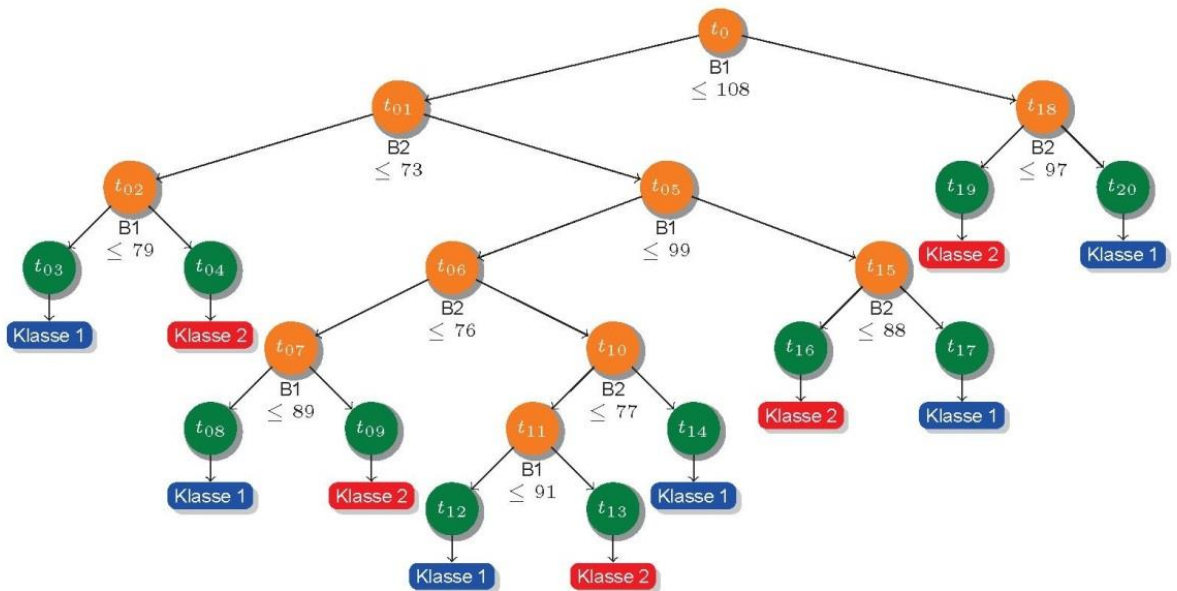


Рис. 3.1. Просте дерево рішень банку даних з двома змінними та двома класами

Суть класифікації за допомогою Random Forest полягає у створенні k дерев

рішень, які потім використовуються у моделі класифікації. Цей процес, відомий як навчання алгоритму, включає в себе використання випадкових вибірок даних з набору даних у пропорціях, визначених користувачем. Більшість дерев рішень генерують первинний результат, який відображає розподіл даних. Окремо, 30% даних, що не включені у формування дерев рішень, можуть бути використані для попередньої кількісної оцінки результатів класифікації. Кожне дерево рішень вміщує статистичні дані, що допомагають визначити розподіл класів усередині набору даних.

Ключовим аспектом у побудові дерева рішень є вибір оптимальної точки для розділення тренувальних даних на «чистіші» підмножини. Функція, яка визначає чистоту цього розділення або наявність домішок, має відповідати декільком критеріям:

1. Максимум при рівномірному розподілі класів: Функція повинна показувати високі значення, коли класи розподілені рівномірно.
2. Мінімум для «чистих» класів: Функція повинна показувати низькі значення, коли дані відносяться до одного класу.
3. Симетрія щодо ймовірності класів: Функція повинна гарантувати однакову ймовірність для обох класів.

Простою функцією, що задовольняє ці вимоги, є індекс Джині. Ймовірність того, що спостереження в точці (x) належатиме до класу (c) , у випадку багатьох класів, можна визначити за допомогою наступної формули:

$$i(t) = 1 - \sum_k p^2(k|t)$$

Дуда та ін. наводять інший часто використовуваний приклад визначення чистоти ентропії:

$$i(t) = - \sum_k p(k|t) \log p(k|t)$$

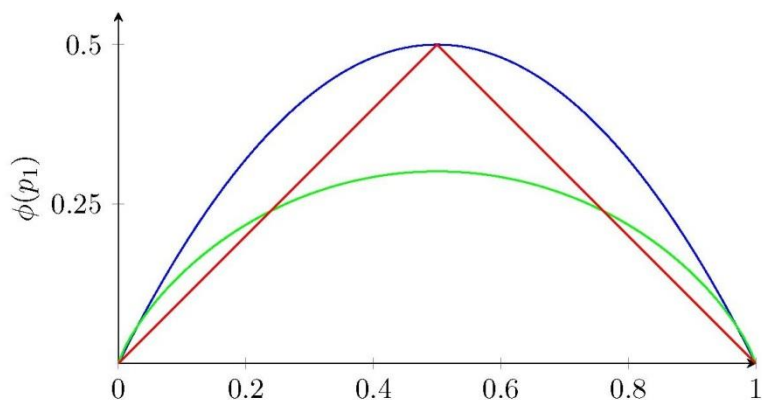
Іншою простою функцією для визначення наявності домішок, є функція чистоти помилки відсутності класифікації:

$$r(t) = 1 - \max p(k|t)$$

Брайман та інші дослідники визнають, що індекс Джині, хоча й відповідає вищевказаним критеріям чистоти, не є найбільш оптимальною функцією для розподілу. Однією з причин є те, що вона може приводити до формування численних точок розподілу на одній і тій же змінній, що може ускладнити процес прийняття рішень та збільшити складність моделі.

На графіках, представлених на рисунку 1.11, можна побачити поведінку функції чистоти для двох класів. Ці зображення демонструють, що всі три розглянуті функції чистоти задовольняють зазначені вимоги: вони досягають максимальних значень, коли класи розподілені рівномірно, і мінімальних значень для ситуацій, коли класи "чисті". Крім того, ці функції є симетричними відносно ймовірностей класів, що забезпечує баланс між категоріями при прийнятті рішень.

Така поведінка є критичною для ефективної класифікації, але важливо враховувати потенційні недоліки функції, як-от тенденція до надмірного розбиття. У практичних ситуаціях, де потрібно балансувати між точністю та складністю моделі, може виявитися корисним розгляд інших функцій чистоти або додаткових методів регуляризації для оптимізації процесу побудови дерев рішень у рамках алгоритму Random Forest.



Ймовірність p_1

— Індекс Джині; — Ентропія; — Помилка класифікації

Рис. 3.2. Функції чистоти класифікації у випадку двох класів

За Брайманом та ін. апіорна ймовірність того, що спостереження відповідатиме певному класу, може бути визначена або подана як відношення N_k/N , де N_k – загальна кількість спостережень в класі k , N , – загальна кількість спостережень у тренувальних (навчальних) даних Lk . Кількість спостережень класу в точці k позначено як $N_k(t)$.

Ймовірність того, що спостереження відповідатиме класу та випадє в вузол, визначається за формулою [24]:

$$i(t) = 1 - \sum_k p^2(k|t)$$

Ймовірність того, що випадкова точка потрапить саме у вузол t , можна визначити за формулою:

$$p(t) = \sum_k p(k, t)$$

На основі вище описаних формул ймовірність того, що точка t у вузлі відповідатиме класу, може бути визначена за наступною формулою:

$$p(k|t) = p(k, t)/p(t)$$

Сума $p(k|t)$ дорівнює 1.

Якщо $\pi(k)$ прирівняти до N_k/N , то $p(k|t)$, то можна розрахувати з $N_k(t)/N_k$ як відношення спостережень класу k та спостережень у вузлі t .

Відношення спостережень для дочірніх вузлів t_L та t_R можна вивести з наступних формул:

$$p_L = \frac{p(t_L)}{p(t)}, \quad p_R = \frac{p(t_R)}{p(t)}$$

при чому $p_L + p_R = 1$.

На основі домішок у вузлі t та підвузлів t_L та t_R , а також зі співвідношень спостережень впливає показник якості отриманої точки розподілу δ у вузлі на основі чистоти:

$$\Delta(\delta, t) = i(t) - p_L i(t_L) - p_R i(t_R)$$

Оптимальне відокремлення даних у деревах рішень (CART) досягається, коли різниця між домішками у батьківському вузлі та зваженими домішками у дочірніх вузлах є максимальною. Це означає, що точка розподілу вибирається так, щоб максимізувати чистоту обох виходячих підмножин. Використовуючи цю ідею, можна визначити найкращу точку розподілу для кожної змінної, яка мінімізує домішки і забезпечує чіткіші класифікації.

CART також дає можливість врахувати некласифіковані дані, що збільшує гнучкість і точність алгоритму в реальних сценаріях, де деякі дані можуть бути недоступними або втраченими. Це робить CART особливо корисним у випадках, коли деякі тренувальні дані відсутні або є погано організованими.

Однією з ключових переваг дерев рішень є їхня простота та інтерпретованість. Їх легко зрозуміти, що робить їх ідеальними для швидкого аналізу та розуміння структури даних. Вони також досить надійні, навіть якщо деякі тренувальні дані змінені, що робить їх стійкими до невеликих змін у вхідних даних.

Втім, як зазначають Дуда та ін., одновимірне визначення точки розподілу, хоч і спрощує оптимізаційне завдання, може призводити до певних обмежень. Зокрема, класи, які перетинаються в одному вимірі, але можуть бути розділені в більш високовимірному просторі, не завжди будуть чітко відокремлені. Це може призвести до необхідності багатьох точок розподілу для адекватного розділення класів, особливо у випадку більш складних даних.

Таким чином, хоча CART є потужним інструментом для класифікації та регресії, важливо враховувати його обмеження та за потреби розглядати альтернативні або допоміжні методи для підвищення точності аналізу.

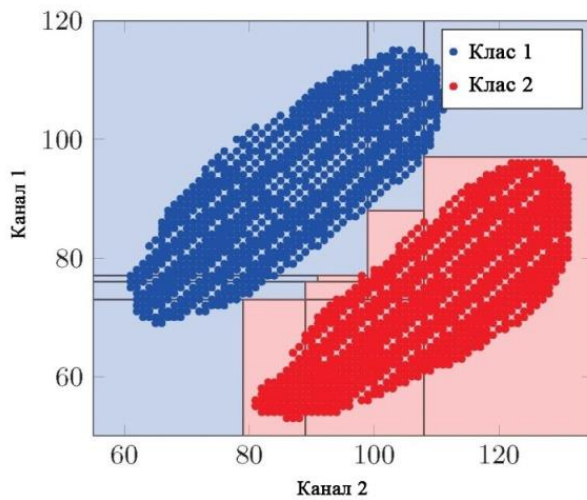


Рис. 3.3. Два відокремлювані класи, які можуть бути розподілені лише при наявності множини точок розподілу (відповідне дерево рішень представлено на рис. 3.1)

Під час формування дерев рішень, можна оцінити попередню точність, з якою вхідні дані розділяються на класи. Це дає уявлення про те, наскільки добре модель може класифікувати нові, невідомі дані. На рисунку 1.13, можна побачити, що точність класифікації даних на два класи може досягати до 98%, що демонструє високу ефективність дерев рішень у розділенні даних. Особливо, це стосується сценаріїв, де кількість дерев варіюється від 30 до 300.

Після завершення процесу формування дерев, створюється файл з розширенням .rfs. Цей файл містить всі необхідні параметри моделі, яка була вивчена на основі навчальних даних. Завдяки цьому файлу можна здійснити класифікацію нових даних за допомогою алгоритму Random Forest. Використання цього методу дозволяє не лише ефективно класифікувати дані, а й забезпечити можливість повторного використання моделі, збереженої в .rfs файлі, для подальших досліджень чи операцій.

Таким чином, Random Forest демонструє високу точність і гнучкість у класифікації даних, роблячи його потужним інструментом в багатьох сферах аналізу даних, включаючи дистанційне зондування, медичні дослідження, фінансовий аналіз та багато інших.

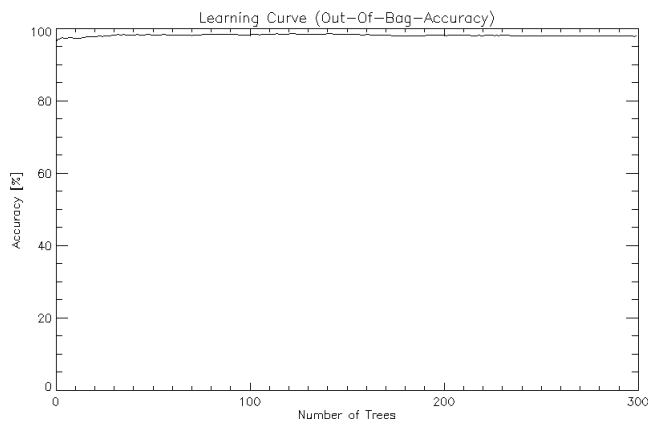


Рис. 3.4. Параметризація вхідних даних для формування дерев рішень

Для оцінки ефективності та якості класифікації за допомогою алгоритму RandomForest використовуються різноманітні статистичні показники. Це дозволяє не тільки визначити загальну точність моделі, але й зрозуміти її поведінку та можливі обмеження.

Показник загальної точності є одним з основних індикаторів якості класифікації. Він визначає відсоток випадків, коли класифікація була виконана правильно. Цей показник надає загальне уявлення про ефективність моделі, але не завжди дає повну картину, оскільки не враховує структуру помилок.

Каппа-коефіцієнт - це більш вдосконалений індикатор, який порівнює точність класифікації з тим, що могло б бути отримано випадковим чином. Це дає уявлення про те, наскільки результати класифікації є значущими та надійними. Коефіцієнт може варіюватися від 0 (жодного відповідності з довідковими даними) до 1 (повна відповідність), і чим вище його значення, тим краще.

Матриця змішування (confusion matrix) - це дуже важлива інструментальна таблиця, яка показує, як різні класи були класифіковані. Вона дозволяє виявити, для яких класів модель працює добре, а для яких - є проблеми. Зокрема, вона може виявити випадки, коли певні класи систематично плутаються з іншими, що може вказувати на необхідність подальшої оптимізації моделі.

Всі ці індикатори інтегровані у програмне забезпечення ENVI, що робить процес оцінки моделі зручним та ефективним. Використання цих показників дозволяє не тільки оцінити загальну ефективність класифікації, але й забезпечити глибше

розуміння її поведінки, виявити потенційні проблеми та напрямки для подальшого вдосконалення.

3.2. Показники часових рядів матеріалів дистанційних спостережень

Часові ряди даних можуть базуватись на рядах сирих цифрових значень, коефіцієнтів відбиття (%), чи на показниках, обчислених на основі оригінальних даних залежно від аналізу. Такими показниками можуть бути геофізичні показники, індексні показники, тематичні показники, топографічні показники та текстурні показники (Kuenzer et al. 2014). Геофізичні показники визначаються для фізичних одиниць, окремі випадки обчислення показників відбиття верхніх шарів атмосфери, відбиття поверхні, фотосинтетично активна радіація, температура земної поверхні, індекс листової поверхні. На противагу, індексні показники є безрозмірними, такі як (NDVI), the enhanced vegetation index (EVI), the Leaf Area Index (LAI), the soil water index (SWI) та багато інших індексів чи компонентів спектрального простору, таких як компоненти Landsat Tasseled Cap.

Обидва геофізичні показники та індексні показники можуть бути використані для довгострокових часових рядів, які можуть бути аналізовані з кроком і день, тиждень, місяць, рік як середнє значення, максимальне чи мінімальне значення, стандартне відхилення, варіація та тренд.

Дуже часто використовуваний набір даних індексів для аналізу часових рядів є набір даних GIMMS обчислених із знімків AVHRR і наявних на сайті NOAA. GIMMS є скороченням від “Global Inventory Monitoring and Modelling Studies. Просторова роздільність таких даних є 4 км, такі дані є наявними із 1981. Подібними часовими рядами рослинних індексів є глобальні продукти S-10, 10-ти денний синтез знімків SPOT-VEGETATION (VGT-S10), які містять дані NDVI та всі спектральні канали просторовою роздільною здатністю 1 км від 1998 року. Ці дані є у вільному доступі через 3 місяці після їх попаданні в архів VEGETATION. Крім того, синтезований із знімків MODIS набір даних (MODIS13) який наявний в збірнику 5, містить індекси NDVI та EVI із просторовою роздільною здатністю 250 м та 16-ти денним інтервалом із 1999 (Terra) та 2002 (Aqua).

Тематичні показники є показники, отримані для аналізу часових серій на основі

на класифікації чи регресійних методів. Типово, набір даних із тематичних змінних є бінарним набором даних з двома класами таких як «ліс-неліс». Деякі наявні продукти дозволяють проводити аналіз часових рядів на певних тематичних шарах, таких як глобальний шар снігового покриву (Dietz et al. 2012, 2013, 2014) та глобальний шар водної поверхні (Klein et al. 2015), обидва розроблені групою динаміки наземного покриву, німецького центру дистанційних даних, німецького космічного агентства (DLR). Іншою тематичною змінною може бути фракція типу поверхні, отримана від спектрального неміксованого аналізу рангів від 0 до 100% і показують як багато пікселів покриваються відповідною фракцією, та часові ряди фракцій зображення дозволяють візуалізацію змін яких субпіксельних компонент (Kuenzer et al. 2008).

Топографічні змінні включають висоту, експозицію, стрімкість, розчленованість. Проте, зазвичай, використовується тільки висота над рівнем моря для аналізу часових рядів. Текsturні змінні зараз використовуються не зовсім часто в аналізі часових рядів, однак сучасна методи часової та просторової сегментації та зростання комп'ютерних потужностей призводить до зростання відповідних змінних.

3.4. Компоненти часових рядів матеріалів дистанційних методів спостережень.

Кожен довгостроковий часовий ряд звичайно складається із трьох компонентів. По-перше, довгостроковий тренд напрямку, по-друге, сезонний, систематичний рух, по-третє, додатковий нерегулярна, безсистемна, короткотермінова флуктуація. Аналіз часових рядів супутникових знімків націлений на визначення динаміки наземного покриву може використовувати кожен із трьох компонентів. Залежно від напрямку досліджень, це може бути сезонний, короткотерміновий, чи трендовий компонент на який є спеціальний інтерес – чи всі три разом (Verbesselt et al. 2010). Коли дослідники клімату можуть мати спеціальний інтерес на довгострокові тренди (температура земної поверхні, тривалість снігового покриву, зміна рівня світового океану, і т.п.), коротко термінова флуктуація є часто відповідною на короткотермінові зміни в окультурених натуральних ресурсах (хвороби рослин, пожежі, природні катастрофи і т.п.). Такі короткотермінові флуктуації які часто називають залишком - є залишкові компоненти якщо тренд і сезонні компоненти забираються з часових

рядів. Слід також зазначити, що цей залишковий компонент не містить ніякого шуму в даних. Часто застосовуваний інструмент для виділення компонентів часових рядів є інструмент BFAST tool (Breaks For Additive Season and Trend), котрий був розроблений Verbesselt et al. (2010). BFAST дозволяє розподіл часових рядів на компоненти тренду, сезонні компоненти та залишкові компоненти разом з методами для виявлення та оцінки змін в часових рядах.

3.3. Близький до реального часу моніторинг

Основне припущення підходів моніторингу NRT, представлених у цьому розділі, полягає в тому, що нормальна часова траєкторія даної спектральної смуги протягом року може бути зафіксована моделлю часових рядів. Наприклад, фенологічний цикл, який зазвичай зустрічається в лісах, можна наблизити сумою тригонометричних функцій різної частоти та компонента тренду або поліноміальних траєкторій. Отже, нещодавно доступні спостереження можна порівняти з прогнозом на основі моделі. Потім значні зміни спектральної ознаки, можливо пов'язані з порушенням лісу, вказуються на статистично значущі відхилення між наявними спостереженнями та прогнозом. Взагалі на дуже низьку частку лісів впливають збурення, і тому основні зусилля для підходу до моніторингу NRT полягають у точній ідентифікації людей, що пережили час, у часових рядах. Це вимагає точного попереднього оброблення часових рядів. Для попередньої обробки доступні операційні програмні засоби, такі як програмне забезпечення Sen2Cor, що надається ESA, однак ідентифікація хмар та тіней хмари все ще є слабкою. Тому в рамках дослідження було докладено значних зусиль для створення вдосконалених ланцюгів попередньої обробки 22, що покращує виявлення хмарних і хмарних тіней, однак все-таки не вдалося реалізувати повністю оптимальне рішення. Іншим аспектом є геометрична точність спільної реєстрації часового ряду Sentinel-2. ESA планує застосувати автоматизовану геометричну спільну реєстрацію часового ряду Sentinel-2 до глобального еталонного зображення. Тому конкретних розробок для здійснення цього кроку попередньої обробки в поточному ланцюзі обробки не проводилося. Ми очікуємо значного підвищення точності виявлення за допомогою спільно зареєстрованих часових рядів. Демонстрація підходу NRT в тестовій зоні Сколівські Бескиди в прикордонному

регіоні між Україною та Польщею показує потенціал для майже в реальному часі виявлення основних порушень, таких як пов'язані з пошкодженнями короїда, стрес від посухи, відмирання лісу і спричинені людиною лісові порушення. Головною перевагою траєкторійного підходу є те, що фіксується дата отримання зображення, щодо якої вперше спостерігалось порушення. Це дозволяє оцінити динаміку порушень, наприклад, динаміка стадій зараження короїдом.

Поточні місії спостереження Землі, що використовують космічні оптичні датчики, набувають величезного обсягу даних. Завдяки високоякісній геореференції та радіометричному калібруванню супутникових знімків можна створити щільний часовий ряд вимірюваних значень сірого кольору для будь-якого заданого спектрального каналу на рівні пікселів. Основне припущення підходів моніторингу близького до реального часу (NRT), представлених у цьому розділі, полягає в тому, що нормальна часова траєкторія даної спектральної смуги протягом року може бути зафіксована моделлю часових рядів. Наприклад, фенологічний цикл, який зазвичай зустрічається в лісах, можна наблизити сумою тригонометричних функцій різної частоти та складовою тренду. Отже, нещодавно доступні спостереження можна порівняти з прогнозом на основі моделі. Значні зміни спектральної ознаки, котрі можливо пов'язані з порушенням лісу, відображаються статистично значущими відхиленнями між наявними спостереженнями та прогнозом.

Більшість алгоритмів, розроблених для моніторингу NRT, стикаються з проблемами через природний характер вхідних даних. Стійкий хмарний покрив спричиняє великі розриви у часових рядах, що, у свою чергу, ускладнює чітке відображення та відстеження внутрішньорічної сезонної динаміки. Недосконалість у процесі хмарного скринінгу призводить до не маскованих хмар та хмарних тіней, залишених у часовому ряду. Надійний метод виявлення змін повинен включати механізм, щоб уникнути плутанини між сторонніми людьми, викликаними хмарами, та реальними змінами. Існує кілька встановлених методів, заснованих на траєкторіях, здатних вирішити ці проблеми. Серед них - монітор BFAST ("Перерви для сезону та тенденції добавок", Verbesselt et al. 2012), CCDC (Виявлення та класифікація постійних змін, Чжу та ін., 2014) або метод, запропонований Brooks et al. (2014). Усі

алгоритми, згадані в попередньому пункті, були застосовані до даних Landsat. Одне з дослідницьких питань, на яке звертається дане дослідження, тому стосується застосованості траєкторійних підходів до даних Sentinel-2. Теоретично збільшена геометрична роздільна здатність та щільність часових рядів повинні привести до кращих результатів. Однак місія Sentinel-2 ще знаходиться на ранній стадії, особливо щодо якості попередньої обробки, порівняно з програмою Landsat (Рис 3.5) У цьому дослідженні досліджується застосованість державних просторових моделей та фільтрації Кальмана, а також кластерний підхід з метою врахування просторово-часового характеру даних.

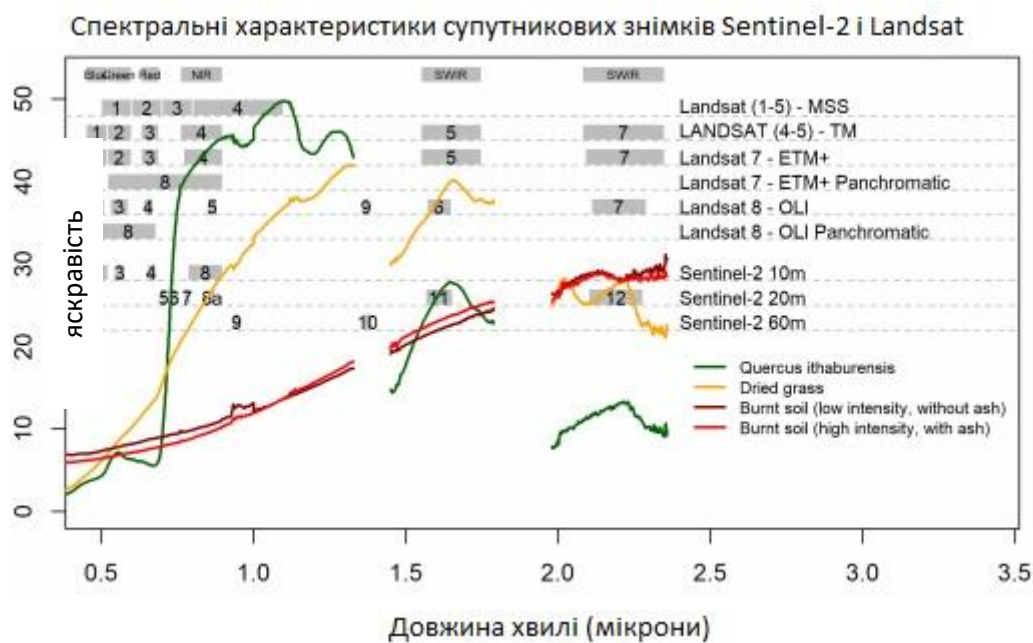


Рис. 3.5 - Спектральні канали всіх сенсорів Landsat та Sentinel-2

Супутникові знімки Landsat-8 OLI та Sentinel-2 мають багато подібного, як показано на рисунку.

РОЗДІЛ 4. ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ

4.1 Застосування і основні принципи роботи Google Earth Engine

Google Earth Engine є потужною платформою для аналізу геопросторових даних, що надається Google. Вона дозволяє дослідникам, науковцям та розробникам аналізувати величезні масиви даних про земну поверхню і спостереження за навколишнім середовищем. Ось основні принципи та застосування Google Earth Engine:

1. **Величезна база даних:** Google Earth Engine забезпечує доступ до декадної історії супутникових знімків Землі та інших геопросторових даних, що оновлюються регулярно. Це включає дані з таких супутників, як Landsat, Sentinel та інші.

2. **Масштабний аналіз:** Платформа дозволяє обробляти великі обсяги даних в хмарі, не вимагаючи від користувачів потужної обчислювальної техніки. Користувачі можуть виконувати складні аналітичні запити на великих масштабах.

3. **Інтерактивність та доступність:** Google Earth Engine має веб-базований інтерфейс, що дозволяє користувачам легко візуалізувати результати, проводити аналіз та ділитися своїми проектами. Код можна писати прямо в браузері за допомогою JavaScript або Python API.

4. **Спільнота та співпраця:** Платформа підтримує спільноту користувачів, які діляться своїми скриптами, алгоритмами та результатами. Це стимулює співпрацю та інновації у галузі аналізу геопросторових даних.

5. **Застосування:** Google Earth Engine використовується в широкому спектрі областей, включаючи моніторинг зміни клімату, агрономію, лісове господарство, водне господарство, урбаністику та інші. Він може допомогти в ідентифікації тенденцій, моделюванні сценаріїв, та прийнятті обґрунтованих рішень на основі даних.

Google Earth Engine продовжує розвиватися, забезпечуючи більш потужні інструменти та ресурси для аналізу геопросторових даних, що робить його незамінним інструментом для дослідників і професіоналів у різних галузях.

4.2 Python і Java Script та їх застосування в Google Earth Engine

Google Earth Engine (GEE) підтримує дві основні мови програмування: Python і JavaScript, кожна з яких має свої особливості та сценарії застосування.

JavaScript у Google Earth Engine:

1. Веб-базований Code Editor: GEE має вбудований веб-редактор коду для JavaScript, що робить його легкодоступним та зручним для швидкого старту. Це ідеально для тих, хто хоче швидко протестувати ідеї або навчитися основам без встановлення додаткового ПЗ.

2. Візуалізація та інтерактивність: JavaScript-інтерфейс дозволяє легко візуалізувати дані та результати аналізу на карті, працювати з шарами та взаємодіяти з користувачем через графічний інтерфейс.

3. Спільнота: Оскільки це була перша підтримувана мова GEE, існує велика кількість прикладів коду та скриптів, розроблених спільнотою, які можна використовувати як основу для власних проектів.

Python у Google Earth Engine:

1. Python API: GEE надає API для Python, що дозволяє використовувати всі можливості GEE у звичному для багатьох розробників Python-середовищі. Це зокрема корисно для інтеграції з іншими науковими та аналітичними бібліотеками Python.

2. Гнучкість та розширюваність: Python є мовою загального призначення з великим співтовариством та багатим набором бібліотек для обробки даних, машинного навчання, веб-розробки тощо. Це робить його ідеальним для комплексних проектів, де аналіз GEE є лише частиною загального рішення.

3. Скриптинг та автоматизація: Python ідеально підходить для написання скриптів, які можуть автоматизувати рутинні завдання, обробляти великі набори даних та інтегруватися з іншими системами.

Сценарії застосування:

- **Навчальні проекти та прототипування:** JavaScript у веб-редакторі GEE чудово підходить для освітніх цілей, швидкого прототипування та невеликих проектів.

- **Комплексний аналіз даних:** Python ідеально підходить для роботи з комплексними проектами, які вимагають інтеграції з іншими аналітичними інструментами та системами.

- **Візуалізація та інтерактивність:** Для проектів, що вимагають високої ступеня візуалізації та інтерактивності, JavaScript у веб-редакторі GEE може бути більш зручним.

Обираючи між Python і JavaScript для роботи з Google Earth Engine, важливо враховувати особисті уподобання, вимоги проекту та наявний досвід. Обидві мови забезпечують потужні можливості для аналізу геопросторових даних і можуть бути використані для різноманітних завдань, від освітніх проектів до комплексних наукових досліджень.

Для розробки карт лісових насаджень було визначено три основні категорії згідно з дослідницькими даними: території з домінуванням хвойних порід, території з перевагою листяних порід, та ділянки без лісового покриву. Втім, на початковій стадії класифікації виявилось, що деякі пікселі водночас помилково кваліфіковані як хвойні породи через схожість спектральних характеристик. Для коректування цієї проблеми було створено додаткові сигнатури для водних об'єктів на основі візуальної інтерпретації космічних знімків.

У дослідженні використовувались зображення з мультиспектральних сенсорів Landsat TM за червень та листопад, а також радіолокаційні зображення ERS-2 і їх комбінації. З метою збільшення масиву даних для класифікації, проведено сегментацію існуючих зображень за допомогою методу попиксельної контурної сегментації (Superpixel Contour Segmentation, SPC).

Сегментація здійснювалася як ітеративний процес, що дозволяє виділити однорідні сегменти зображення, які відповідають реальним об'єктам на поверхні Землі. Виділення сегментів базувалося на аналізі статистичного розподілу пікселів в межах кожного сегменту, з використанням ітераційної сегментації (maximum-a-

posteriori (MAP)). Параметри G і B у методі SPc дозволяли регулювати масштаб та форму сегментів відповідно. Оптимальні значення цих параметрів були визначені експериментально в рамках спільного дослідницького проекту.

Для аналізу обрали зображення з різними значеннями параметра G (10, 15, 20 та 25). Класифікація включала аналіз різноманітних зображень, в тому числі радіолокаційних ERS-2, мультиспектральних Landsat TM за різні місяці, а також їх комбінацій з сегментованими зображеннями.

Точність класифікації оцінювалась за такими параметрами як перша та друга класова точність, загальна точність та індекс Каппа, використовуючи 30% тренувального набору даних для верифікації. Це дозволило виявити оптимальні параметри та синтезовані зображення, які забезпечують найкращі результати класифікації.

Водночас, візуальний аналіз дешифрованих зображень підтвердив, що пікселі водних поверхонь часто помилково класифікуються як хвойні насадження, що вимагало додаткової корекції та розподілу цих ділянок на відповідні класи порід.

Результати попередньої оцінки точності та відповідні висновки були представлені у таблиці 4.1.

Таблиця 4.1. – Результати попередньої оцінки точності класифікації зображень

Зображення	Хвойні дерев.		Листяні дерев.		Загальна точність	Індекс Каппа
	I*	II	I	II		
ERS	92	77	65	86	80,43	0,59
L5_June	96	89	80	92	90,21	0,79
L5_November	95	90	82	91	90,79	0,80
L5_JuneNov_15	95	94	88	91	93,19	0,85
L5_JuneNov_20	95	88	79	90	89,32	0,77
L5_JuneNov_25	95	95	91	90	94,21	0,87
ERS+L5_JN_10	97	89	80	95	91,24	0,81
ERS+L5_JN_15	96	99	98	91	96,64	0,92

ERS+L5_JN_20	96	93	87	92	93,16	0,85
ERS+L5_June_10	97	97	94	93	96,39	0,92
ERS+L5_Nov_10	97	87	79	95	90,37	0,79
ERS+L5_June_15	97	94	89	94	94,56	0,88
ERS+L5_Nov_15	97	99	97	94	97,54	0,94
ERS+L5_June_20	92	95	89	84	91,65	0,81
ERS+L5_Nov_20	95	93	87	90	92,68	0,83

Примітка: I – перша класова точність, II – друга класова точність.

Як випливає з даних таблиці 4.1, попередня оцінка точності дешифрування перевищує 80%, що демонструє придатність алгоритму Random Forest для аналізу рослинності. Зазначено, що радіолокаційний знімок показав найменшу точність, однак її значно можна підвищити шляхом комбінації з мультиспектральними зображеннями. Особливо вражаючі результати були досягнуті для зображень, які включали сегментовані фрагменти (в таблиці 4.1 вони виділені жирним шрифтом). Найбільш ефективною виявилася комбінація радіолокаційних зображень ERS-2, мультиспектральних зображень Landsat TM (знятих у листопаді 2010 року), та сегментованих зображень із параметром G=15, що забезпечує високий контраст між листопадковими і хвойними породами.

На основі отриманих даних обрали лише три зображення з найвищими показниками попередньої оцінки точності для подальшої детальної класифікації. Використовуючи сформовані моделі параметризації для відповідних синтезованих зображень, проведено класифікацію за допомогою алгоритму Random Forest в програмному продукті ENVI. Результати класифікації представлені на рисунку 4.1.

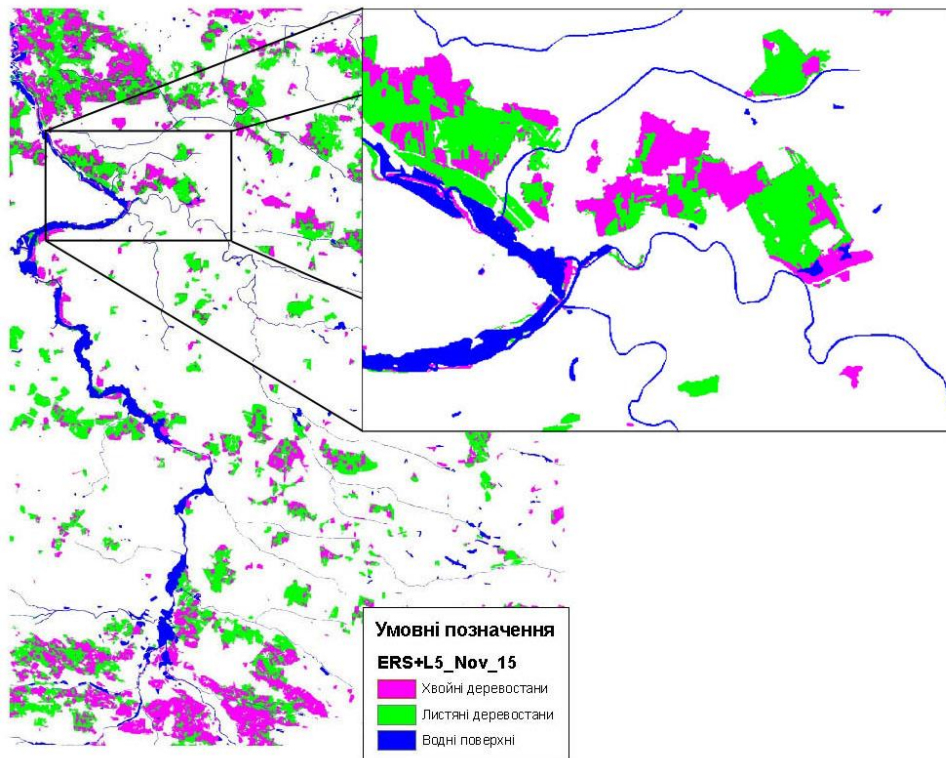


Рис. 4.1. Класифіковане синтезоване зображення (радіолокаційний знімок, мультиспектральний знімок, отриманий у листопаді 2010 р. та сегментоване зображення з $G=15$) за алгоритмом Random Forest

З рисунку 4.1 стає зрозуміло, що окрім хвойних та листяних лісів, було також виявлено клас водних територій. Це стало можливим завдяки фотоінтерпретації та цифровій обробці мультиспектральних знімків Landsat TM та векторних даних ESRI. Розпізнавання рослинності за допомогою алгоритму Random Forest демонструє ефективність у забезпеченні актуальної картографічної та статистичної інформації, а також управління процесом вибору даних для статистичного аналізу. Попередні результати показують, що точність класифікації перевищує 80%, що підтверджує ефективність використання Random Forest для розпізнавання типів рослинності. Синтезовані зображення, які поєднують радіолокаційні та мультиспектральні дані, зокрема з додаванням сегментованих фрагментів, значно підвищують точність розпізнавання порівняно з однотипними знімками.

Виділено три зображення з найвищою точністю для подальшого детального аналізу. Класифікація цих зображень в ENVI за допомогою Random Forest демонструє

класифіковані ділянки на рис. 4.1. Окрім того, для підвищення достовірності отриманих результатів, проведено порівняння із даними лісовпорядкування на великій площі, що допомогло виявити подібності та можливість використання будь-якого з обраних зображень для подальшого аналізу. Перевірка точності через побудову матриці помилок вказала на добру точність аналізу, незважаючи на використання недосконалих даних лісовпорядкування. Таким чином, матеріали польових досліджень можуть служити для попередньої оцінки, а лісовпорядкування - для детального аналізу динаміки лісового покриву.

Класифіковані карти створені для двох часових відтинків 2021 і 2021 рр. Порівнювались та визначались місця втрати лісового покриву. Для прикладу використані карти на територію «Сколівських Бескид». На таких картах відображені лісництва квартали і виділа, котрі позначаються відповідними індексами .

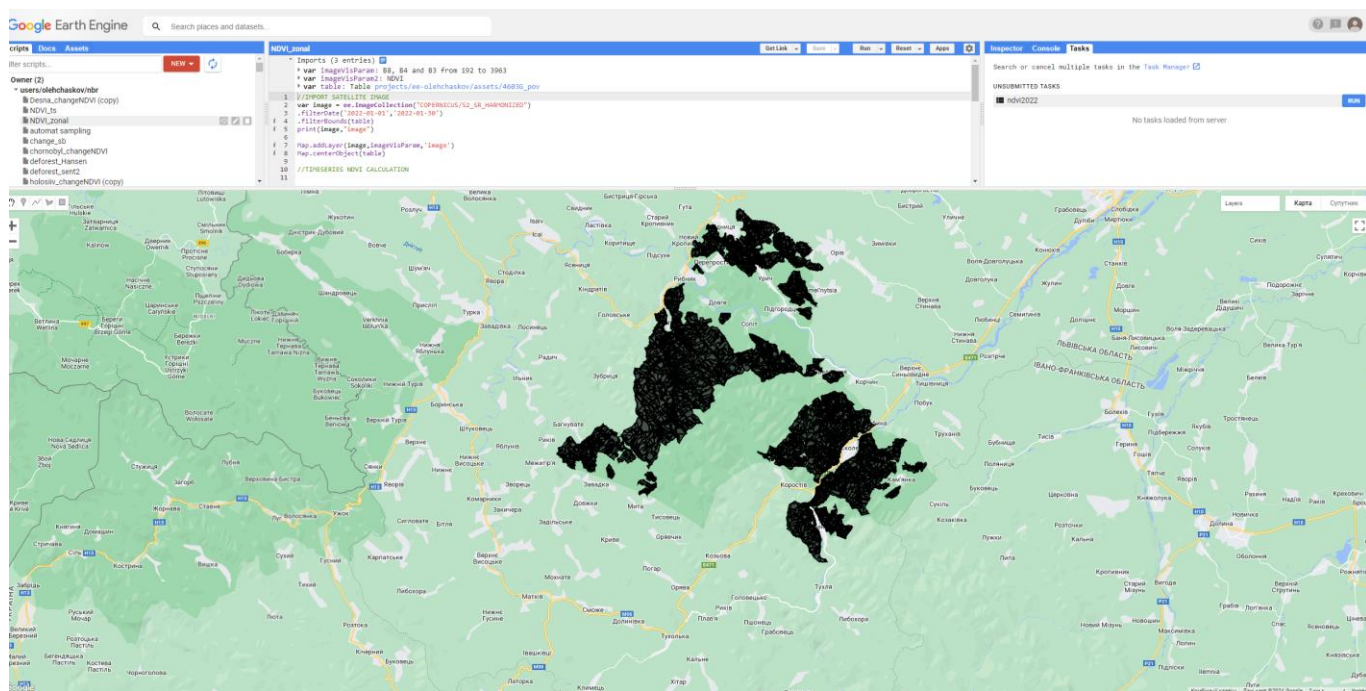


Рис. 4.2. Територія Сколівських Бескид із суміщеними лісовими картами.

Такі карти містять картографічну і описову інформацію про окремі частини лісових насаджень(виділи), а також квартали та лісництва. Така інформація поряд з точними координатами розміщення є достатньою для прийняття рішень відповідними державними органами.

4.3. Google monitoring для генерування повідомлень

Даний розділ описує механізм формування та надсилання повідомлень на електронну пошту (e-mail) про настання певних подій, а саме змін у ландшафті (лісовому покриві) на певних ділянках. Система розпізнає зміни у лісовому покриві ділянок Сколівських Бескидів і відсилає заздалегідь визначене значення метрики у Google Monitoring (Stackdriver). Назва метрики відповідає певній ділянці. Програмний код, написаний на Python, авторизується за допомогою [OAuth2](#) та системного облікового запису (service account) у Google Cloud та використовуючи Google API надсилає метрику і її значення у Google Monitoring (Stackdriver). Python код може працювати фактично у будь-якому середовищі, який має доступ в інтернет та підтримує Python 3.8 або вищу версію. Тестування відбувалось на віртуальній машині у Google Cloud (Linux Ubuntu), Google Cloud shell та локальний комп'ютер з ОС Windows 10. Також можливе використання serverless Cloud functions.

Розглянемо детальніше як і що саме потрібно налаштувати у Google Cloud для роботи нашої Python програми.

1. Для проекту у Google Cloud потрібно підключити необхідні API: [Cloud Monitoring API](#), [Cloud Channel API](#), [Error Reporting API](#), [Google Earth Engine API](#).
2. Для створення необхідних налаштувань у Google Monitoring потрібно надіслати будь-яке значення нашої метрики, наприклад метрика з назвою test_tree1 із значенням 28

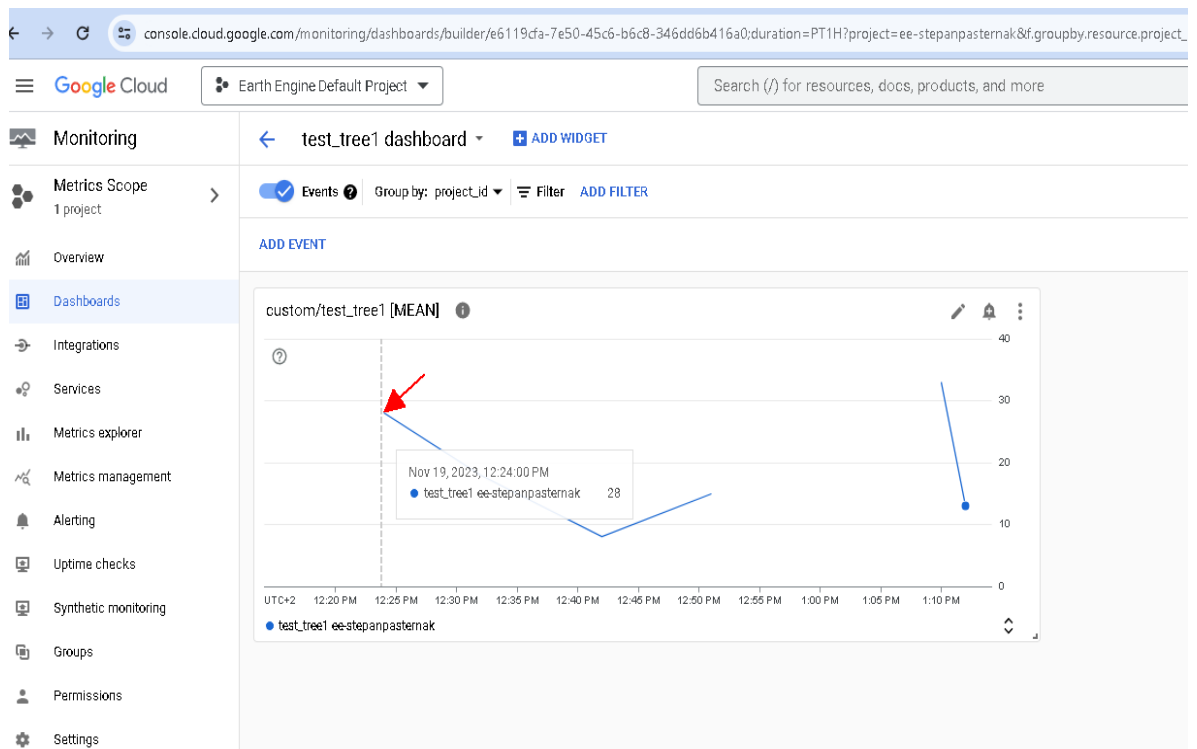


Рис. 15 - Відображення метрики test_tree1 на Metrics Dashboard

Далі потрібно створити канал відсилання повідомлень (notification channel). Окрім повідомлень на електронну пошту, їх можна відсилати на мобільні пристрої (SMS повідомлення), Google Chat, PagerDuty, Slack, Webhooks (Endpoint URL) та Pub/Sub.

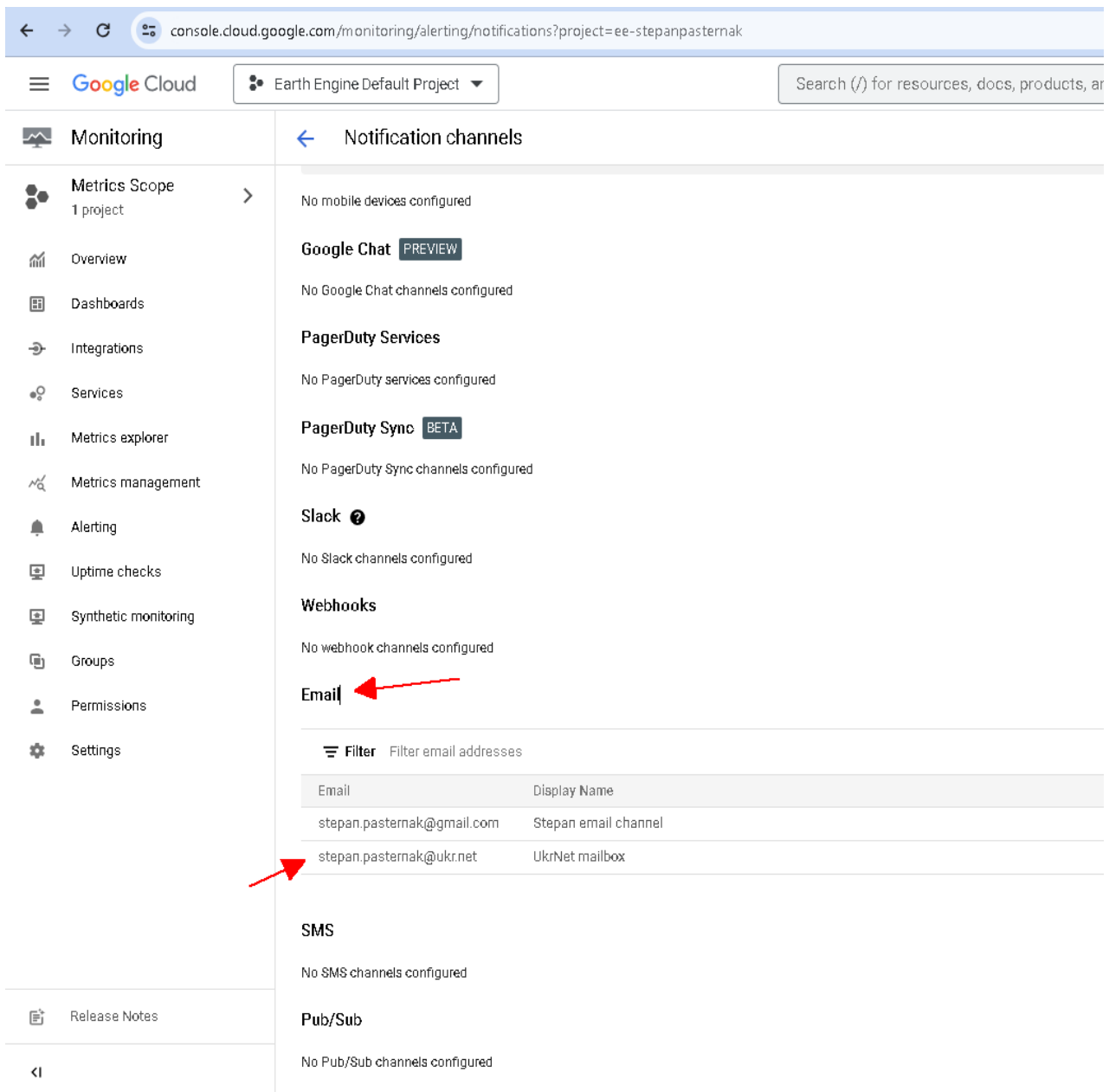


Рис. 16 – Налаштування каналу відсилання повідомлень на електронну пошту

3. Налаштування політик алертів (Alerting policies). Цей розділ налаштувань є найважливішим і поєднує у собі кілька важливих елементів, а саме:

3.1. Вибір конкретної метрики та умови (condition)

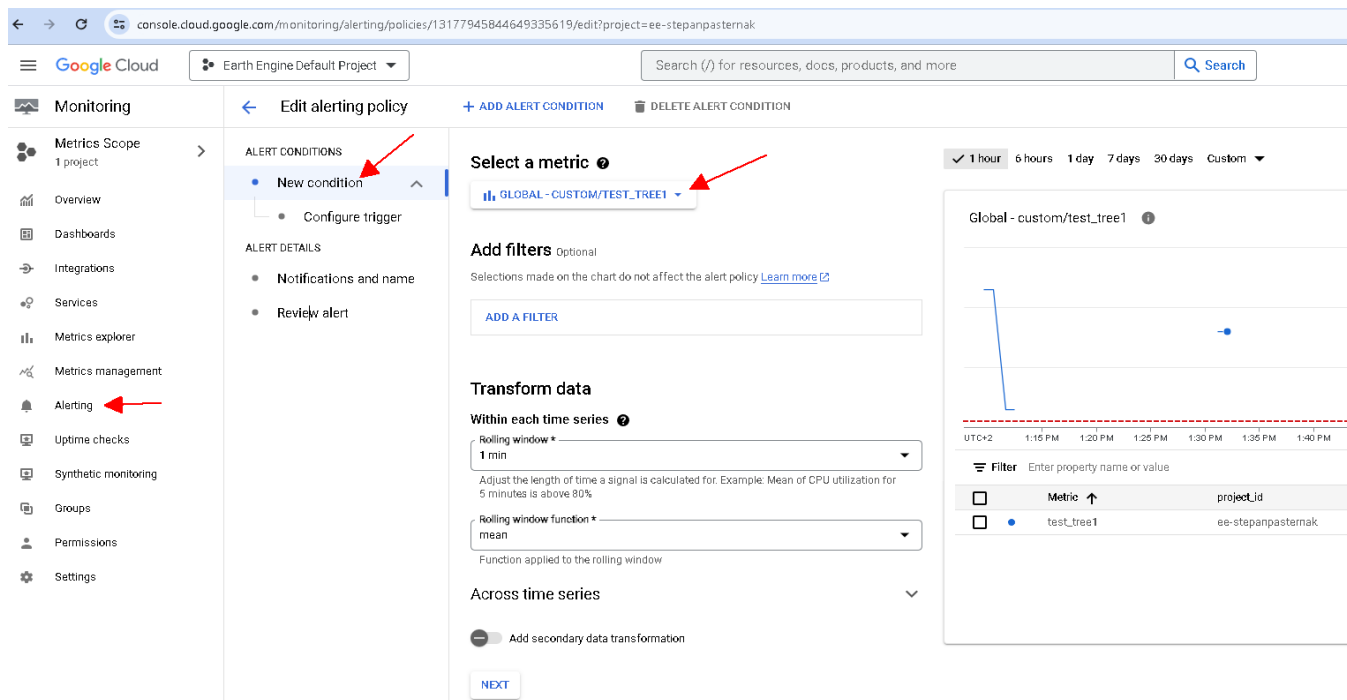


Рис. 17 – Налаштування умов та метрики для політики алерту

3.2. Конфігурація спрацювання алерту (trigger)

У даній роботі були використані наступні параметри спрацювання алерту (trigger)

- Тип умови (Condition type) - поріг спрацювання (threshold)
- Значення порогу спрацювання (threshold position) – перевищення (above threshold)

• Значення порогу (threshold value)

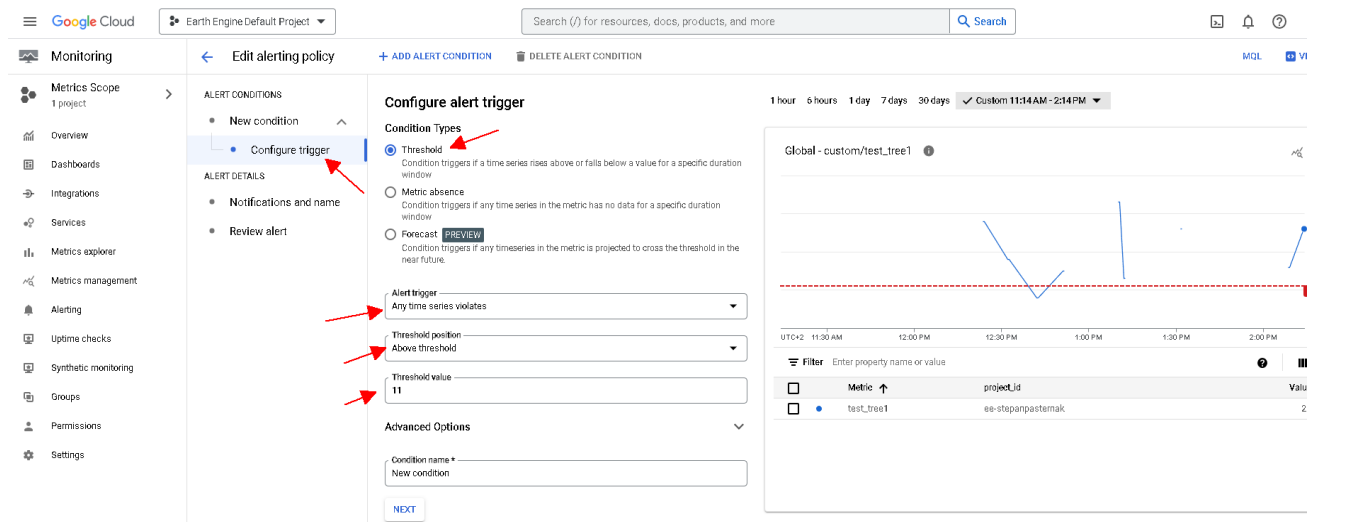


Рис. 18 – Налаштування порогу спрацювання для політики алерту

3.3. Налаштування повідомлень

Даний пункт містить налаштування повідомлень при спрацюванні алерта, а саме:

- Канал відсилання повідомлень (Notification Channel) – фактично вибір одної або кількох адрес електронної пошти куди буде надіслано повідомлення.
- Тема електронного повідомлення (Notification subject line)
- Текст самого повідомлення (Documentation). Підтримуються деякі елементи HTML, а також є своя мова розмітки.

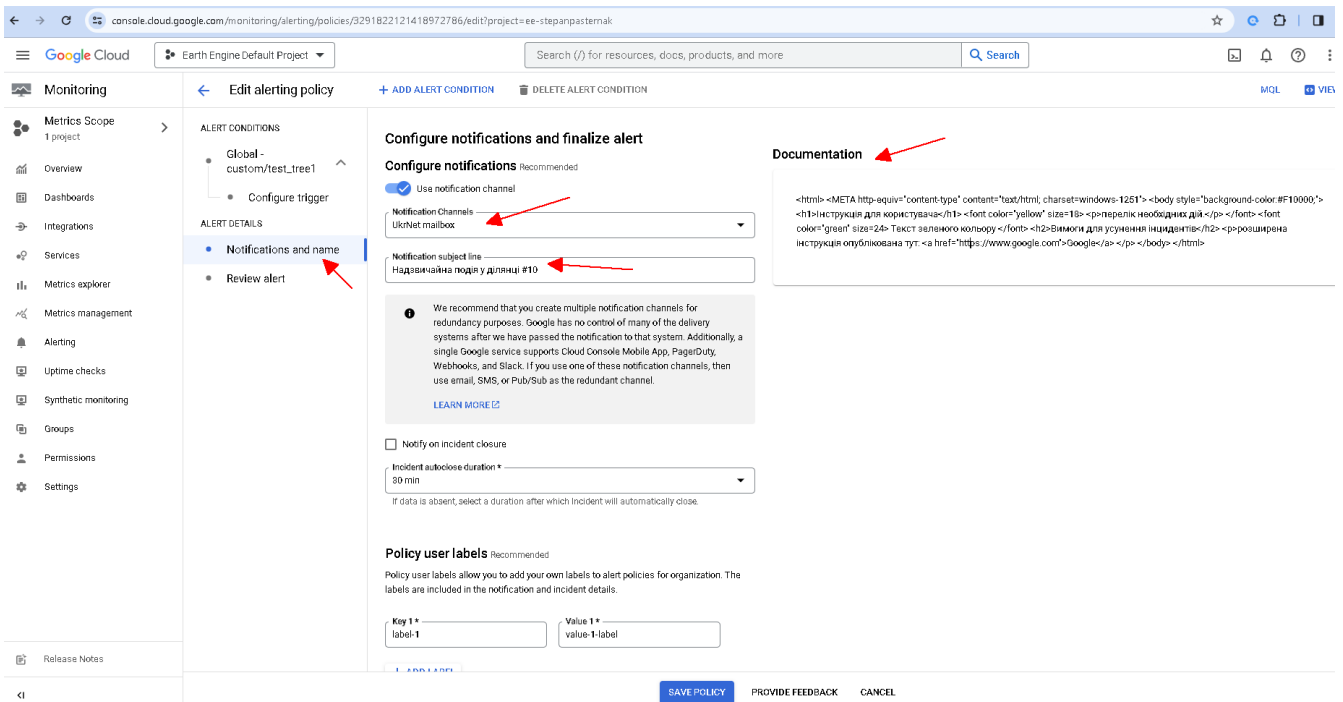



Рис. 19 – Налаштування повідомлень алерту.

Повідомлення на електронну пошту будуть мати вигляд:

[ALERT] Надзвичайна подія у ділянці #10 ←

Google Cloud Alerting <alerting-noreply@google.com> сьогодні, 15:24 ☆

Кому: stepan.pasternak@ukr.net ←

 VIEW INCIDENT

Alert firing

Global - custom/test_tree1

custom/test_tree1 for ee-stepanpasternak is above the threshold of 20.000 with a value of 47.000.

Summary

Start time
Nov 19, 2023 at 1:24PM UTC (less than 1 sec ago)

Project
[ee-stepanpasternak](#)

Policy
[test_alert_policy2](#) ←

Condition
Global - custom/test_tree1

Metric
custom.googleapis.com/test_tree1

Threshold
above 20 ←

Observed
47.000

Resource labels
project_id: ee-stepanpasternak

Policy labels
label-1: value-1-label

Policy documentation ←

Інструкція для користувача

перелік необхідних дій.

Текст зеленого кольору

Вимоги для усунення інцидентів

розширена інструкція опублікована тут: [Google](#)

Висновок

Описане вище рішення для відсилання повідомлень при виявленні надзвичайної ситуації на певній ділянці лісового покриття максимально використовує засоби Google Cloud без застосування будь яких сторонніх програмних засобів чи розробок.

Також використана опція налаштувань закриття алерту після 30 хвилин без додаткових повідомлень так як нас цікавить саме настання події а не відновлення

попередніх умов.

4.4. Особливості побудови системи попереджень про втрати лісового покриву на території Львівщини.

Алерт-системи попереджень про втрати лісового покриву — це спеціалізовані інструменти моніторингу, які використовуються для виявлення та повідомлення про зміни в лісових масивах. Вони засновані на аналізі даних з супутників, інших дистанційних джерел, а також на місцевих спостереженнях. Ось деякі ключові особливості та принципи роботи цих систем:

1. Джерела даних:

- **Супутникові знімки:** Більшість систем використовують дані з супутників, таких як Landsat, Sentinel та інші, щоб регулярно отримувати оновлення про стан лісів.

- **Дрони та повітряні фотографії:** У деяких випадках для детальнішого моніторингу використовуються дрони або літаки.

- **Місцеві спостереження:** Інформація від місцевих спільнот і науковців може також інтегруватися в систему для підвищення точності.

2. Технології обробки даних:

- **Алгоритми обробки зображень:** Сучасні алгоритми дозволяють аналізувати зміни в лісовому покриві, порівнюючи знімки різних періодів.

- **Машинне навчання:** Для більш точного розпізнавання змін використовуються методи машинного навчання та штучного інтелекту.

- **Геоінформаційні системи (ГІС):** Для візуалізації та аналізу просторових даних використовуються ГІС.

3. Системи попередження:

- **Автоматичні сповіщення:** Системи можуть автоматично надсилати сповіщення вченим, органам управління та іншим зацікавленим сторонам.

- **Інтерактивні платформи:** Часто існують веб-платформи, де можна переглядати актуальну інформацію, карти змін і отримувати детальні звіти.

4. Застосування:

- **Моніторинг вирубки лісів:** Швидке виявлення незаконної вирубки та інших втрат лісового покриву.

- **Оцінка впливу стихійних лих:** Слідкування за наслідками пожеж, ураганів та інших природних катастроф.

- **Планування землекористування:** Допомога урядам і організаціям у прийнятті обґрунтованих рішень щодо використання та захисту лісових територій.

5. Виклики:

- **Чутливість до умов:** Хмарність, тіні та інші фактори можуть ускладнювати інтерпретацію даних.

- **Потреба в оновленні:** Для точності необхідні регулярні оновлення даних.

- **Точність та помилки:** Не завжди легко розрізнити природні зміни від антропогенних, існує ризик помилкових тривог.

Як тестовий регіон використано територію Сколівських Бескид (рис. 4).

Системи попередження про втрати лісового покриву відіграють важливу роль у збереженні лісів та біорізноманіття, надаючи своєчасну та цінну інформацію для реагування на потенційні проблеми.

РОЗДІЛ 5. РОЗРОБЛЕННЯ СТАРТАП-ПРОЕКТУ

5.1 Опис ідеї проекту

Ідея стартової системи оповіщення про вирубку лісів на основі супутникових зображень передбачає використання супутникових технологій для моніторингу та виявлення змін у лісовому покриві, надаючи сповіщення в режимі реального або майже в реальному часі відповідним зацікавленим сторонам. Ця система спрямована на вирішення критичної проблеми вирубки лісів шляхом раннього виявлення та реагування на незаконні рубки, розчищення земель та інші дії, які сприяють втраті лісів.

Ось схема того, як може працювати така система сповіщень про порушення:

Технологія опрацювання супутникового зображення:

Використовують супутникові зображення високої роздільної здатності з частими повторними переглядами, щоб отримати детальну та актуальну інформацію про лісисті території.

Використовують передові технології дистанційного зондування, такі як радар із синтетичною апертурою (SAR) або багатоспектральне зображення, щоб покращити можливості виявлення.

Автоматичний аналіз зображень:

Запровадження алгоритмів машинного навчання та методів комп'ютерного зору для автоматичного аналізу супутникових зображень.

Навчання системи розпізнавати закономірності, пов'язані з діяльністю з вирубки лісів, як-от видалення деревного покриву, зміни щільності рослинності або поява лісових доріг.

Моніторинг в реальному часі:

Постійно спостереження за великими лісовими масивами в режимі реального часу або з мінімальною затримкою.

Налаштуйте систему, яка може постійно обробляти та аналізувати вхідні супутникові дані, дозволяючи негайно виявляти зміни.

Генерація сповіщень:

Встановлення порогу для виявлення підозрілих або аномальних змін у

ландшафті.

Коли система виявляє потенційну діяльність з вирубки лісів, генерує автоматичні сповіщення, які надсилаються відповідним органам влади, екологічним організаціям та іншим зацікавленим сторонам.

Інтерфейс користувача та інформаційна панель:

Розроблений зручний інтерфейс або інформаційну панель, яка відображає проаналізовані супутникові дані, виділяючи проблемні області.

Надані інтерактивні карти та детальна інформацію про виявлені зміни, що дозволить користувачам провести подальше дослідження та вжити відповідних заходів.

Інтеграція з іншими системами:

Інтеграція системи оповіщення з існуючими платформами екологічного моніторингу, державними базами даних і правоохоронними системами для скоординованої реакції.

Співпрацювати з урядовими та неурядовими організаціями для обміну даними та забезпечення комплексного підходу до боротьби з вирубкою лісів.

Така нова система оповіщення сприятиме глобальним зусиллям по боротьбі з вирубкою лісів, забезпечуючи проактивне втручання та сприяючи своєчасним заходам збереження. Це узгоджується зі зростаючою важливістю технологій та інновацій у вирішенні екологічних проблем.

5.2 Розроблення ринкової стратегії

Створення надійної маркетингової стратегії для системи оповіщення про вирубку лісів на основі супутникових зображень передбачає націлювання на ключові зацікавлені сторони, підвищення обізнаності та встановлення ціннісної пропозиції рішення. Ось комплексна маркетингова стратегія:

1. Визначте цільову аудиторію:

- Визначте та визначте пріоритети ключових цільових аудиторій, включаючи екологічні організації, державні установи, неурядові організації, органи управління лісовим господарством та компанії, які зобов'язуються сталого розвитку.

2. Створіть переконливу ціннісну пропозицію:

- Чітко сформулюйте унікальні переваги та особливості супутникової системи оповіщення про вирубку лісів.

- Підкресліть здатність системи забезпечувати моніторинг у реальному часі, раннє виявлення та розуміння на основі даних для ефективного запобігання вирубці лісів.

3. Створіть онлайн-присутність:

- Розробіть професійний та інформативний веб-сайт, який висвітлює технологію, її можливості та позитивний вплив, який вона може мати на боротьбу з вирубкою лісів.

- Використовуйте методи пошукової оптимізації (SEO), щоб гарантувати, що веб-сайт займатиме високі позиції в релевантних пошукових запитах.

4. Контент-маркетинг:

- Створюйте високоякісний вміст, як-от дописи в блогах, статті та інфографіку, щоб розповісти вашій цільовій аудиторії про проблеми вирубки лісів і важливість проактивного моніторингу.

- Поділіться історіями успіху, тематичними дослідженнями та оновленнями про те, як система сприяла запобіганню вирубці лісів у певних регіонах.

5. Залучення до соціальних мереж:

- Використовуйте платформи соціальних медіа, щоб охопити ширшу аудиторію та створити спільноту навколо збереження довкілля.

- Діліться візуально привабливим вмістом, включаючи супутникові зображення, відео та інфографіку, щоб проілюструвати вплив вирубки лісів і ефективність вашого рішення.

6. Партнерство та співпраця:

- Співпрацюйте з екологічними організаціями, державними установами та міжнародними організаціями, щоб продемонструвати довіру та надійність вашого рішення.

- Отримайте підтримку від відомих екологів або впливових людей у сфері сталого розвитку.

7. Освітні вебінари та майстер-класи:

- Проводьте вебінари та семінари, щоб ознайомити потенційних користувачів із технологією, її застосуванням і найкращими методами запобігання вирубці лісів.

- Взаємодіяти з аудиторією, відповідати на запитання та демонструвати можливості системи.

8. Прес-релізи та ЗМІ:

- Публікуйте прес-релізи, щоб повідомити про ключові етапи, партнерства або успішні заходи, які сприяють системі оповіщення про вирубку лісів.

- Зверніться до відповідних ЗМІ, щоб забезпечити висвітлення та підвищити видимість.

9. Участь у конференціях та заходах:

- Відвідайте галузеві конференції, екологічні саміти та заходи, щоб спілкуватися з ключовими особами, які приймають рішення, демонструвати технології та бути в курсі галузевих тенденцій.

- Розгляньте можливість розміщення стенду або участі в панельних дискусіях для встановлення лідерства в думках.

10. Безкоштовні пробні версії та демонстрації:

- Пропонуйте безкоштовні пробні версії або демонстрації системи оповіщення потенційним користувачам, дозволяючи їм випробувати її можливості з перших вуст.

- Збирайте відгуки та відгуки, щоб завоювати довіру та вирішити будь-які проблеми.

11. Урядова та регуляторна адвокація:

- Взаємодіяти з державними органами та регуляторами, щоб відстоювати впровадження передових технологій моніторингу в зусиллях зі збереження навколишнього середовища.

- Продемонструйте, як система узгоджується з існуючою екологічною політикою та ініціативами та підтримує їх.

12. Постійна підтримка клієнтів:

- Забезпечення чудової підтримки клієнтів для вирішення запитів, допомоги з адаптацією та забезпечення плавної інтеграції системи сповіщень у робочі процеси

користувачів.

Об'єднавши ці елементи в комплексну маркетингову стратегію, ви зможете ефективно повідомити про цінність супутникової системи оповіщення про вирубку лісів, завоювати довіру та сприяти прийняттю серед ключових зацікавлених сторін, які віддані справі збереження навколишнього середовища.

5.3. Вимоги до технічного та програмного забезпечення

Створення супутникової системи оповіщення про вирубку лісів вимагає поєднання технічних і програмних компонентів для збору, обробки, аналізу та поширення інформації. Ось основні задіяні компоненти:

Джерело даних із супутникових зображень:

- Постачальники супутникових даних: такі як Махар, Planet Labs, , щоб отримати доступ до найновіших супутникових зображень високої роздільної здатності.

Завантаження та обробка даних:

- Система прийому даних супутникових зображень у режимі реального часу або через регулярні проміжки часу.

- Хмарне сховище (наприклад, AWS S3, Google Cloud Storage) для ефективного зберігання великих обсягів даних супутникових зображень.

Автоматичний аналіз зображення:

- Алгоритми машинного навчання та методи комп'ютерного зору для аналізу супутникових зображень на предмет вирубки лісів.

- Frameworks глибокого навчання такі як Google Earth Engine для навчання та розгортання моделей глибокого навчання.

Система моніторингу в реальному часі:

- Потокова обробка даних для обробки даних у реальному часі.
- База даних для зберігання та отримання даних моніторингу в реальному часі.

Генерація сповіщень:

- Системи, засновані на правилах для визначення критеріїв для ідентифікації

потенційної діяльності з вирубки лісів.

- Запуск подій для створення сповіщень, коли виконуються заздалегідь визначені критерії.

Інтерфейс користувача та інформаційна панель:

- Розробка інтерфейсу використовуючи веб-технології (наприклад, React, Angular, Vue.js) для візуалізації супутникових зображень, місць сповіщень і додаткової інформації.

- Картографічні служби такі як Google Maps або Mapbox для інтерактивного відображення карт.

Інтеграція із зовнішніми системами:

- API: розробляйте API для полегшення інтеграції із зовнішніми системами, такими як урядові бази даних, платформи екологічного моніторингу та системи правоохоронних органів.

- Веб-хуки: Впроваджуйте веб-хуки для спілкування в реальному часі та обміну даними із зовнішніми системами.

Інтеграція ГІС (геоінформаційна система):

- Програмне забезпечення ГІС для аналізу просторових даних і функцій картографування.

- Геопросторові бібліотеки (наприклад, GDAL, Fiona) для обробки геопросторових даних.

Ретельно відбираючи та інтегруючи ці технічні та програмні компоненти, ваш стартап може створити комплексну та ефективну супутникову систему оповіщення про вирубку лісів. Крім того, бути в курсі нових технологій і супутникового прогресу буде мати вирішальне значення для підвищення можливостей системи з часом.

ВИСНОВКИ

В ході роботи було створено застосунок який дозволяє на основі точкових зйомок контурів лісосіки з'єднати їх в єдиний полігон, підготувати геометрію полігону для обчислення просторових показників, обчислити показники та внести їх у стандартну таблицю і сформувати план лісосіки із стандартним його представленням. Для побудови плагіна використані можливості QGIS та написано програмний код мовою Python.

При створенні застосунку для високоточних ГПС та їх використанні в лісовому господарстві можна зробити кілька важливих висновків:

- використання високоточних ГПС дозволяє підвищити ефективність лісового господарства через точніше планування та виконання різноманітних лісогосподарських робіт;

- застосунок допомагає забезпечити високу точність та надійність геодезичних даних, що робить їх придатними для прийняття обґрунтованих рішень;

- використання ГПС сприяє ефективному моніторингу та управлінню лісовими ресурсами, допомагаючи зберегти їх сталість та здоров'я;

- застосунок допомагає покращити лісове планування, забезпечуючи точні та оновлені дані для прийняття стратегічних рішень;

- з використанням високоточних ГПС можливо зменшити витрати та ризики, пов'язані з непередбаченими лісозаготівельними та іншими роботами;

- застосунок має зручний та інтуїтивний інтерфейс, що сприяє легкому використанню та прийняттю користувачами;

- застосунок вигідно впливає на розвиток сталого лісового господарства, забезпечуючи збалансований підхід до використання лісових ресурсів.

У цілому, успішна реалізація застосунку для високоточних ГПС в лісовому господарстві сприяє покращенню ефективності, сталості та цифровій трансформації галузі.

- Успішний застосунок повинен мати зручний та інтуїтивний інтерфейс, що сприяє легкому використанню та прийняттю користувачами.

- Застосунок вигідно впливає на розвиток сталого лісового господарства, забезпечуючи збалансований підхід до використання лісових ресурсів.

У цілому, успішна реалізація застосунку для високоточних ГПС в лісовому господарстві сприяє покращенню ефективності, сталості та цифровій трансформації галузі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Геопортал «Ліси України» [електронний ресурс]. – Режим доступу : - <http://forestry.org.ua>
2. Де Мерс, Майкл Н. Географические информационные системы: Пер. с англ. – М.: Дата , 1999. – 489с.
3. Миклуш С.І. Геоінформаційні системи в лісовому господарстві / С.І.Миклуш, М.П. Горошко, О.Г. Часковський. – Львів: «Камула», 2007. – 128 с.
4. Морозов В. В. Геоінформаційні системи в агросфері / В. В. Морозов, К. С. Лисогоров, Н. М. Шапоринська. – Херсон. – 223 с.
5. Часковський О., Андрейчук Ю., Ямелинець Т., Застосування ГІС у природоохоронній справі на прикладі відкритої програми QGIS. [Текст] : навч. посіб. / О. Часковський, Ю. Андрейчук, Т.Ямелинець. — Львів : ЛНУ ім. Івана Франка, Вид-во Простір-М, 2021. — 228 с. — ISBN 978-617-7746-79-8.
6. Quantum GIS, Руководство пользователя, Версия 1.6.0 [Текст]’Соріарог’, 2011, 221 с.
7. K. S. Reynolds, M. J. Kelty. Integration of GPS and GIS in Forest Management."Forest Management and Planning"
8. L. S. Evans, R. A. Cunha. "GIS and GPS Technology for Sustainable Forest Resource Management"."Geospatial Technology in Environmental Management"

ДОДАТКИ

Додаток А

Код в Python для об'єднання точок в полігони

```
from qgis.core import QgsProcessingFeatureSourceDefinition, QgsVectorLayer, QgsProject
import processing
from os.path import basename, dirname, splitext, join
from qgis.PyQt.QtCore import *
from qgis.PyQt.QtWidgets import *
from qgis.core import *
from qgis.PyQt import uic
from .p2o_encodings import getEncodings, getDefaultEncoding, setDefaultEncoding
from .p2o_engine import Engine, P2OError
```

```
BASE_FORM_CLASS, _ = uic.loadUiType(join(dirname(__file__), 'frmPoints2One.ui'))
```

```
class points2One(QDialog, BASE_FORM_CLASS):
```

```
    def __init__(self, iface):
```

```
        super(points2One, self).__init__()
```

```
        self.iface = iface
```

```
        self.setupUi(self)
```

```
        self.wInputLayer.setFilters(QgsMapLayerProxyModel.PointLayer)
```

```
        self.wBrowse.clicked.connect(self.outFile)
```

```
        self.wSort1.toggled.connect(self.sort1_toggled)
```

```
        for combo in (self.wGroupField, self.wSortField1, self.wSortField2):
```

```
            combo.setLayer(self.layer())
```

```
            combo.setCurrentIndex(0)
```

```
        self.populate_encodings(getEncodings())
```

```
        self.show()
```

```
    def layer_name(self):
```

```
        """Return the selected input layer name as unicode."""
```

```
        return str(self.wInputLayer.currentText())
```

```
    def layer(self):
```

```
        """Return the selected input layer as a QgsMapLayer instance."""
```

```

return self.wInputLayer.currentLayer()

def output_geometry(self):
    """Return the selected output geometry."""
    if self.wCreateLines.isChecked():
        return QgsWkbTypes.LineString
    else:
        return QgsWkbTypes.Polygon

def close_lines(self):
    """Return whether lines must be closed."""
    return self.wCloseLines.isChecked()

def group(self):
    """Return whether grouping by attribute is enabled."""
    return self.wGroup.isChecked()

def group_field(self):
    """Return the name of the grouping field."""
    if self.group():
        return str(self.wGroupField.currentText())

def sort_fields(self):
    """Return the names of sorting fields as a list."""
    fields = []
    if self.wSort1.isChecked():
        fields.append(str(self.wSortField1.currentText()))
    if self.wSort2.isChecked():
        fields.append(str(self.wSortField2.currentText()))
    return fields

def output_encoding(self):
    """Return the output encoding as unicode."""
    return str(self.wEncoding.currentText())

```

```

def check_input(self):
    """Check whether the input is valid, raise if not."""
    layer = self.layer()
    if layer is None:
        msg = self.tr('Please select an input layer')
        raise P2OError(msg)

    if self.group_field() == "":
        msg = self.tr('Please select a field to group by')
        raise P2OError(msg)

    for field in self.sort_fields():
        if not field:
            msg = self.tr('Please select a field for sorting')
            raise P2OError(msg)

    if not self.output_path():
        msg = self.tr('Please specify output shapefile')
        raise P2OError(msg)

def populate_encodings(self, names):
    """Populate the combo box of available encodings."""
    self.wEncoding.clear()
    self.wEncoding.addItemNames(names)
    index = self.wEncoding.findText(getDefaultEncoding())
    if index == -1:
        index = 0 # Make sure some encoding is selected.
    self.wEncoding.setCurrentIndex(index)

def update_progress_bar(self):
    """Update the progress bar."""
    self.wProgressBar.setValue(self.wProgressBar.value() + 1)

def _accept(self):
    self.check_input()
    layer = self.layer()

```

```

self.wProgressBar.setRange(0, layer.dataProvider().featureCount())
setDefaultEncoding(self.output_encoding())
engine = Engine(
    layer,
    self.output_path(),
    self.output_encoding(),
    self.output_geometry(),
    self.close_lines(),
    self.group_field(),
    self.sort_fields(),
    self.update_progress_bar
)

engine.run()

# Show warning
log_msg = '\n'.join(engine.get_logger())
if log_msg:
    warningBox = QMessageBox(self)
    warningBox.setWindowTitle('Points2One')
    message = self.tr('Output shapefile created')
    warningBox.setText(message)
    message = self.tr('There were some issues, maybe some features could not be created.')
    warningBox.setInformativeText(message)
    warningBox.setDetailedText(log_msg)
    warningBox.setIcon(QMessageBox.Warning)
    warningBox.exec_()

if self.wAddResult.isChecked():
    addShapeToCanvas(str(self.output_path()))
self.wProgressBar.setValue(0)
# Convert multipart to singlepart
algSingle = processing.runAndLoadResults(
    "qgis:multiparttosingleparts",
    {'INPUT': str(self.output_path()), 'OUTPUT': 'TEMPORARY_OUTPUT'}
)

```

```

# Add the converted layer to the canvas
layer_single = QgsProject.instance().mapLayer(algSingle['OUTPUT'])
QgsProject.instance().addMapLayer(layer_single)

self.wProgressBar.setValue(0)
def accept(self):
    try:
        self._accept()
    except P2OError as e:
        QMessageBox.critical(self, 'Points2One', str(e))

def sort1_toggled(self, checked):
    if not checked:
        self.wSort2.setChecked(False)

def outFile(self):
    """Open a file save dialog and set the output file path."""
    outFilePath = saveDialog(self)
    if not outFilePath:
        return
    self.setOutFilePath(outFilePath)

def output_path(self):
    """Return the output file path."""
    return self.wOutputFileName.text()

def setOutFilePath(self, outFilePath):
    """Set the output file path."""
    self.wOutputFileName.setText(outFilePath)

def saveDialog(parent):
    """Shows a save file dialog and return the selected file path."""
    settings = QSettings()
    key = '/UI/lastShapefileDir'
    outDir = settings.value(key)

```

```

filter = 'Shapefiles (*.shp)'
outFilePath, _ = QFileDialog.getSaveFileName(parent, parent.tr('Save output shapefile'), outDir, filter)
outFilePath = str(outFilePath)
if outFilePath:
    root, ext = splitext(outFilePath)
    if ext.upper() != '.SHP':
        outFilePath = f'{outFilePath}.shp'
    outDir = dirname(outFilePath)
    settings.setValue(key, outDir)
return outFilePath

```

```

# Convenience function to add a vector layer to canvas based on input
# shapefile path (as string).
def addShapeToCanvas(shapeFilePath):
    layerName = basename(shapeFilePath)
    root, ext = splitext(layerName)
    if ext == '.shp':
        layerName = root
    vlayer_new = QgsVectorLayer(shapeFilePath, layerName, "ogr")
    ret = QgsProject.instance().addMapLayer(vlayer_new)
    return ret

```