



## Харківський національний університет радіоелектроніки

Факультет Інформаційно-аналітичних технологій та менеджменту  
(повна назва)Кафедра Інформатики  
(повна назва)Рівень вищої освіти другий (магістерський)Спеціальність 122 Комп'ютерні науки  
(код і повна назва)Освітня програма Інформатика  
(повна назва освітньої програми)

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_  
(підпис)

«\_\_\_\_\_» \_\_\_\_\_ 20\_\_ р.

**ЗАВДАННЯ**  
НА АТЕСТАЦІЙНУ РОБОТУстудентові Павлова Марія Олександрівна  
(прізвище, ім'я, по батькові)1. Тема роботи Дослідження технологій редагування відео та машинного навчання для застосунок TikTokзатверджена наказом по університету від «23» жовтня 2020 року № 1428 Ст2. Термін подання студентом роботи до екзаменаційної комісії 25 листопада 2020 р.3. Вихідні дані до роботи Відео, оброблене засобами TikTokперелік використуваних засобів:теоретичні відомості про методи відео обробкитеоретичні відомості про алгоритми штучного інтелекту в мобільних додаткахвідеоматеріали та зображеннястатистичні відомості про застосунок TikTok

4. Перелік питань, що потрібно опрацювати в роботі \_\_\_\_\_

1. Аналіз існуючих методів обробки відео.

2. Порівняння засобів для редагування відео на смартфоні.

3. Аналіз машинного навчання та його застосування в соціальних мережах

4. Обґрунтування сучасності технічних засобів застосунок TikTok

5. Створення демонстраційного відеоролика засобами TikTok з використанням алгоритмів машинного навчання

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) *Актуальність проблеми сучасних технологій обробки відео в мобільних додатках, порівняння функціоналу обробки відео мобільних додатків, методи машинного навчання в сучасних застосунках, статистичні дані застосунку TikTok, демонстраційне відео*

6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

### КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	Отримання завдання на атестаційну роботу	23.10.2020	<b>виконано</b>
2	Аналіз завдання, підбір літератури	24.10.20-29.10.20	<b>виконано</b>
3	Аналіз літератури з досліджуваної проблеми	30.10.20-04.11.20	<b>виконано</b>
4	Аналіз алгоритмів та методів обробки даних	05.11.20-09.11.20	<b>виконано</b>
5	Обробка відео засобами TikTok	10.11.20-11.11.20	<b>виконано</b>
6	Аналіз статистичних даних	12.11.20-15.11.20	<b>виконано</b>
7	Оформлення пояснювальної записки	16.11.20-25.12.20	<b>виконано</b>
8	Перевірка на плагіат	27.11.20	<b>виконано</b>
9	Рецензування	30.11.20	<b>виконано</b>
10	Підготовка презентації та доповіді	30.11.20	<b>виконано</b>
11	Занесення роботи в електронний архів	01.12.20	<b>виконано</b>
12	Попередній захист атестаційної роботи	02.12.20	<b>виконано</b>

Дата видачі завдання 23 жовтня 2020 р.

Студент \_\_\_\_\_  
(підпис)

Керівник роботи \_\_\_\_\_ ст. викл. Путятіна О.Є.  
(підпис) (посада, прізвище, ініціали)

## РЕФЕРАТ/ABSTRACT

Пояснювальна записка до атестаційної роботи: 72 с., 2 табл., 28 рис., 2 дод., 46 джерел.

ЗАСТОСУНОК ТІКТОК, РЕДАГУВАННЯ ВІДЕО, НЕЙРОННІ МЕРЕЖІ, АНАЛІЗ ДАНИХ, МАШИННЕ НАВЧАННЯ, ВІРТУАЛЬНА РЕАЛЬНІСТЬ, АГОРИТМИ СИНХРОНІЗАЦІЇ, МОБІЛЬНІ ДОДАТКИ, СОЦІАЛЬНІ МЕРЕЖІ.

Метою дослідження є вивчення можливостей мобільного застосунку ТікТок, аналіз функціональних можливостей мобільних застосунків щодо технологій редагування відео, машинного навчання, інтегрування засобів у сучасних застосунках.

Об'єктом дослідження є мобільний застосунок ТікТок та його функціональні можливості.

Використано доступні фільтри застосунку, функція з інтегрованим штучним інтелектом для автоматичної синхронізації музичного супроводження із відеоматеріалами. Проведено аналіз статистичних даних щодо впровадження застосунку та порівняння його функціональних можливостей із медійними додатками Facebook, Instagram, Snapchat, Telegram.

У результаті роботи створено демонстраційний відеоролик засобами ТікТок з використанням алгоритмів машинного навчання.

TIKTOK APPLICATION, VIDEO EDITING, NEURAL NETWORKS, DATA ANALYSIS, MACHINE LEARNING, VIRTUAL REALITY, SYNCHRONIZATION ALGORITHMS, MOBILE APPLICATIONS, SOCIAL NETWORKS.

The purpose of the study is to study the capabilities of the mobile application TikTok, analysis of the functionality of mobile applications for video editing technologies, machine learning, integration of tools in modern applications.

The object of research is the mobile application TikTok and its functionality.

Available application filters are used, a function with integrated artificial intelligence for automatic synchronization of music with videos. The analysis of statistical data on the implementation of the application and comparison of its functionality with the media applications Facebook, Instagram, Snapchat, Telegram.

As a result, a demo video was created using TikTok using machine learning algorithms.

## ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів .....	6
Вступ.....	7
1 Технології редагування відео.....	8
1.1 Відео обробка .....	8
1.1.1 Математичне представлення відео обробки.....	14
1.1.2 Сучасні формати відео.....	20
1.1.3 Можливості редактора.....	23
1.2 Огляд технічних засобів.....	24
1.3 Постановка задачі дослідження.....	28
2 засоби Машинного навчання .....	30
2.1 Загальні методи машинного навчання.....	30
2.2 Використання технологій у сучасних застосунках .....	41
3 вивчення та застосування функціоналу застосунку TikTok .....	48
3.1 Огляд можливостей застосунку.....	48
3.1.1 Обробка відео .....	48
3.1.2 Застосування штучного інтелекту.....	54
3.2 Порівняння функціоналу з іншими застосунками .....	59
Висновки .....	65
Перелік джерел посилання .....	66
Додаток А.....	71
Додаток Б .....	72

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

FPS – (frame per second) кількість кадрів в секунду

NTSC – (National Television System Committee) відеостандарт в ряді країн Північної та Південної Америки, а також Японії, Кореї та Тайвані

PAL – (Phase Alternating Line) відеостандарт в Європі, Азії та Австралії

SECAM – (SEquential Couleur Avec Memoire) відеостандарт в Європі, Азії та Австралії

Соц. мережі – соціальні мережі

Стікер – не велике додаткове зображення як оновлена версія смайлів

Screenshot – зафіксований стопкадр

НМ – нейронна мережа

AI – (Artificial Intelligence) штучний інтелект

Арбітраж контенту (перезалив) – це перенесення контенту з однієї платформи на іншу, з метою отримання більшої кількості переглядів, лайків, трафіку без створення нового контенту

Лайк – це умовний вираз схвалення матеріалу, користувачеві, фотографії, що виражаються натисненням однієї кнопки

Месенджер – додаток для обміну миттєвими повідомленнями і відео.

Клікбейт – зазвичай, заголовок, що змушує користувача інтернету натиснути на нього

## ВСТУП

Обробка відео є розповсюдженим засобом перетворення звичайного відео у справжній фільм із доповненням у вигляді ефектів. Впровадження функціоналу для відео обробки у мобільні застосунки виводить цю сферу на новий рівень. Основними ознаками є зрозумілість, ефективність, легкість в застосуванні.

Відеоредактори для смартфонів відрізняються від комп'ютерних програм тим, що вони дозволяють не тільки завантажувати вже відзняте відео але й захоплювати його в робочому режимі. Також є різниця у форматі запису та його розширенні. Надається унікальна можливість бачити попередній результат при використанні масок, як додаткової реальності в режимі online. Існує багато готових фільтрів для створення неповторної комбінації ефектів відео, що робить обробку простішою та цікавішою за комп'ютерні програмні обчислення.

Сьогодні в світі відбуваються колосальні зрушення в галузі нейронних мереж, оскільки машинне навчання, штучний інтелект та інші подібні технології інтегруються в мобільні застосунки та супроводжують людей в повсякденні. Більшість алгоритмічних прийомів, що використовуються в підходах глибокого навчання, відомі і доступні вже деякий час. Значна частина останніх інновацій у цій галузі стосується того, щоб зробити ці методи доцільними для реальних застосувань. Це передбачає розробку та впровадження архітектур, які можуть виконувати ці методи, використовуючи розумну кількість ресурсів за розумний проміжок часу.

У сучасних мобільних додатках застосовуються новітні алгоритми машинного навчання та обробки даних, що дозволяє індивідуалізувати контент для кожного користувача за його особливими інтересами.

# 1 ТЕХНОЛОГІЇ РЕДАГУВАННЯ ВІДЕО

## 1.1 Відео обробка

Комп'ютерна обробка відео це процес редагування відео файлів за допомогою комп'ютерних спеціалізованих програм – відеоредакторів. Процес обробки відео описується в три послідовні і взаємопов'язані етапи: захоплення відео, редагування і кінцеве стиснення файлу. Розрізняють лінійний та нелінійний комп'ютерний монтаж.

Лінійний монтаж відбувається частіше в реальному часі. Відео з декількох джерел (програвачів, камер т.д.) Надходить через комутатор на приймач (ефірний транслятор чи записуючий пристрій). Про лінійний монтаж також говорять в разі процесу урізання сцен у відеоматеріалі без порушення їх послідовності. Це технологія минулого, оскільки вимагає наявність магнітофонів для послідовного запису фрагментів з одної стрічки на іншу [1, 2].

Нелінійний монтаж так називається, тому що монтажерам нарешті вдалося уникнути тривалого перемотування стрічки вперед-назад (лінійної) і вони отримали прямий доступ до необхідного фрагменту відразу. Нелінійний монтаж став можливий завдяки появі комп'ютерів. При цьому всі вихідні матеріали з стрічки записуються на жорсткий диск. Це може виявитися тривалим процесом, але він залежить від ступеня компресії вихідного матеріалу і потужності комп'ютера. Проте, вже існують відеокамери, які дозволяють відразу завантажувати вихідний матеріал на цифровий носій або здійснювати монтаж з носія [3].

Кілька років тому з'явилася можливість виконання відеомонтажу та редагування «всередині» комп'ютера. Ця технологія отримала назву нелінійного монтажу, оскільки дозволила операторам пряме звернення до необхідних кадрів або фрагментами відео, записаним на жорсткий диск комп'ютера, тобто дозволила уникати стомлюючого процесу постійної

(лінійної) перемотки стрічки вперед-назад при перегляді чи пошуку цих фрагментів. Оцифровані фрагменти відео перед записом на диск піддаються компресії (як правило, MJPEG) в 3-10 разів, що неминуче призводить до певної втрати якості (тим меншою, чим менше ступінь компресії).

Основні переваги цифрового нелінійного монтажу – це збереження вихідного рівня якості записаних на диск фрагментів при їх копіюванні (незалежно від числа копій), ширші можливості контролю процесу монтажу, автоматична синхронізація відео зі звуком [3].

Стандартна цифрова система побудована у одно-потоківій архітектурі. Це означає, що при прорахунку реально задіяна тільки одна копія вихідного відео (кліпу) у вигляді AVI-файлу. У той же час, якщо потрібно трохи більше, ніж просто розрізати або склеїти наявні фрагменти, то необхідно сформувати і задіяти другу копію цифрового кліпу (принаймні, його частини). Іншими словами, для створення будь-якого мікшерного переходу чи ефекту між двома кліпами (А і В) в оперативній пам'яті комп'ютера необхідно одночасно містити кадри як закінчення кліпу А, так і початку кліпу В, послідовно завантажуючи їх з жорсткого диску, декомпресуючи і виробляючи прорахунок нових кадрів результуючого кліпу, потім здійснюючи зворотний компресію і запис на диск. Цей процес, нерідко званий рендерінгом (rendering), ілюструється наступною схемою (рис. 1.1) [4].

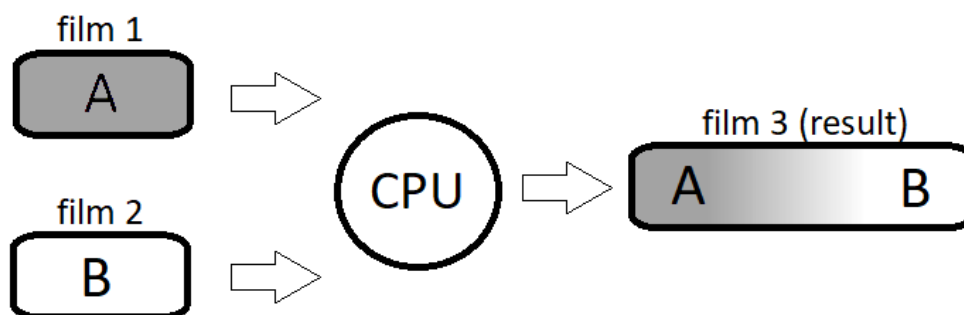


Рисунок 1.1 – Мікшерний перехід між двома кліпами

Подібні обчислення вимагають здійснення мільярдів спеціалізованих операцій над пікселями зображень. Очевидно, що швидкість їх виконання істотно залежить від швидкодії процесора. Стандартні PC виявляються порівняно повільними з точки зору вирішення даного завдання тому що використовують однопотоківі цифрові монтажні системи. Швидкість обробки дуже залежить від продуктивності процесора і складності накладеного ефекту. Актуальним питанням на сьогоднішній день є виконання цифрового монтажу без значних витрат часу на прорахунок. Сучасні технології дозволяють робити первинну обробку в реальному часі коли результат можна відразу бачити на екрані контрольного монітора і / або записати на майстер-стрічку. Це дозволяє використовувати нелінійні монтажні системи не тільки в пост-, а й в оперативному виробництві (тобто online) [4].

Сучасні плати нелінійного монтажу (наприклад, miroVideo DC30plus) для операцій компресії і декомпресії відео ефективно задіють встановлені на них мікросхеми, що, безумовно, прискорює рендеринг, але не призводить до його виконання в реальному часі. Для досягнення останнього необхідно використання спеціалізованого обчислювального пристрою, «заточеного» на прорахунок певного класу ефектів і переходів (таких як, Pinnacle Systems Genie – для тривимірних ефектів). Оскільки набір апаратно виконуються ефектів фіксований для кожного пристрою і залежить від його спеціалізації і моделі, то завжди будуть виникати нестандартні завдання, повністю або частково завантажили процесор комп'ютера.

Одною з переваг цифрового редагування відео є можливість майже необмеженого творчого самовираження, реалізації оригінальних ідей і створення як завгодно складних і неповторних ефектів. Однак навіть наявність подібного спеціалізованого пристрою саме по собі не вирішує проблему рендеринга – на його вхід необхідно одночасно подавати два або більше потоків декомпресованого відео. Загальний рівень розвитку комп'ютерної техніки, досягнутий за останні роки, дозволяє і цю складну

задачу ефективно вирішувати на базі стандартного РС – при певній оптимізації його дискової підсистеми. Таким чином, системи нелінійного монтажу реального часу використовують двохпотоківу плату компресії або декомпресії відео і додаткову плату цифрових ефектів (рис. 1.2). Втім, набір мікросхем для виконання в реальному часі заданих ефектів мікшування може бути встановлений і прямо на платі компресії (наприклад, як у Pinnacle Systems ReelTime – понад 130 двовимірних ефектів виконується в реальному часі). І навіть при цьому може бути використана додаткова плата, що розширює набір апаратно виконуються ефектів (наприклад, Pinnacle Systems ReelTime NITRO [4]).

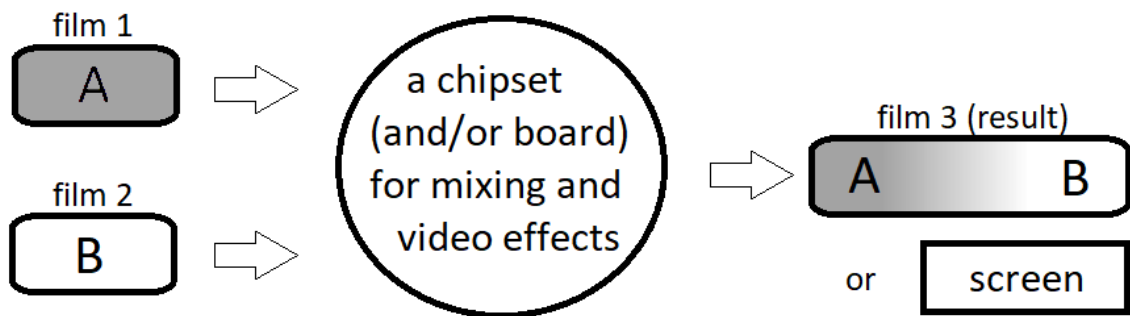


Рисунок 1.2 – Двопотоківий процес монтажу

Оперуючи з двома потоками, подібні цифрові системи можуть виконувати в реальному часі і інші необхідні функції, властиві класичним монтажно-мікшерним аналоговим комплексам, наприклад, титрування (titling) або різні види рип-проекцій («keying», «ключування», проекції з використанням ефектів прозорості). Тобто у дво-потоківих системах титри (так само як і інша комп'ютерна графіка) відразу йде на висновок – минаючи проміжний етап компресії і записи на диск [4].

Відео представляє собою сукупність декількох кадрів, які змінюються за визначеною швидкістю. Частота кадрів – це кількість відображених кадрів за одиницю часу. Загальноприйнята одиниця виміру – кадри в секунду (к/с). Загальносвітовий стандарт частоти кінозйомки і проєкції – 24 кадри в секунду, але на сьогодні маємо 30 fps для телевізійного стандарту NTSC (США, Канада, Мексика, Японія) і 25 fps для стандарту PAL / SECAM (Європа та Азія). Крім того, для зйомок можуть використовуватися як прискорена, так і уповільнена частота кадрів, наприклад, 48 к/с – це частота зйомки і проєкції кінематографічних систем IMAX HD [5].

Людина не здатна відчувати різницю частоти кадрів починаючи з 25 к/с та більше, але до 24 включно різниця стає відчутною. Наприклад, порівняння кроку людини на першому кадрі та другому (рис. 1.3) із частотою запису 1, 10 та 30.

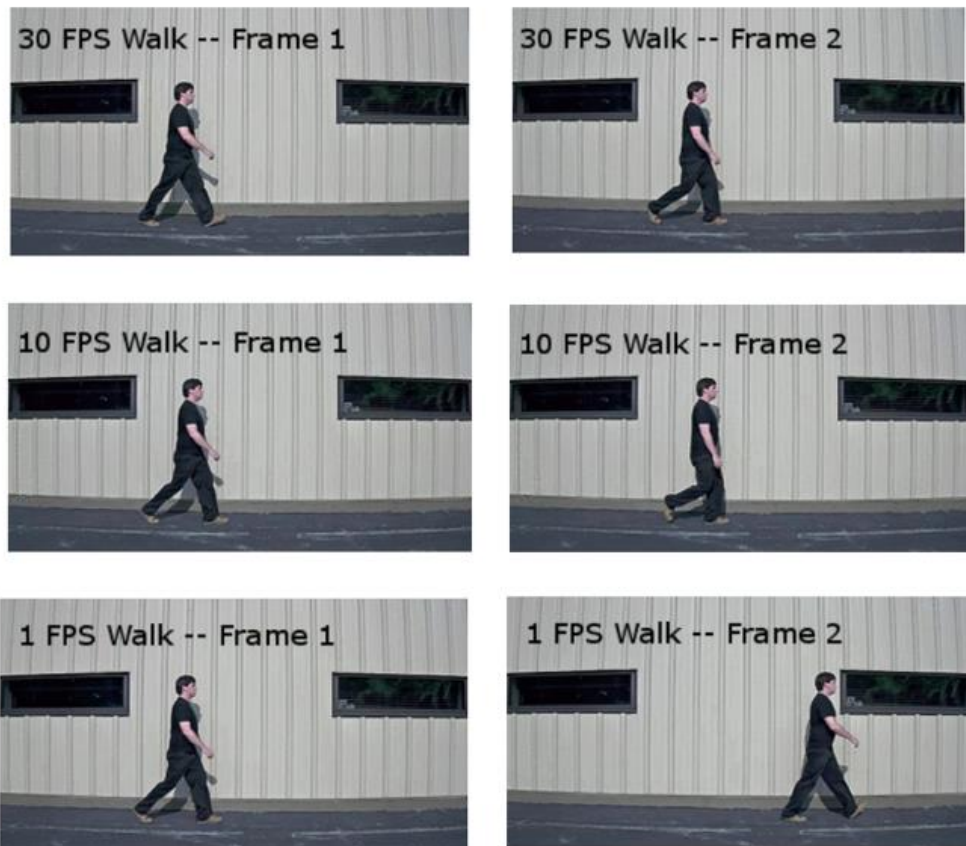


Рисунок 1.3 – Залежність положення людини під час кроку від частоти кадрів, перший кадр (зліва) та другий кадр (справа)

Якщо швидкість руху об'єкта на відео мала, кількість кадрів не сильно впливатиме на передачу інформації. Але при прискоренні руху кількість частота кадрів повинна збільшитись до необхідного для розуміння діапазону. Розглянемо рух людини під час бігу у 1, 10 та 30 кадрах (рис. 1.4) [5].



Рисунок 1.4 – Залежність положення людини під час бігу від частоти кадрів, перший кадр (зліва) та другий кадр (справа)

Від частоти кадрів у секунду також залежить розмір відео ролика. Чим більше значення, тим більше пам'яті необхідно для зберігання запису. Але велика кількість кадрів сприяє плавності руху та підвищенню якості кожного окремого кадру. Якщо кількість кадрів не велика, але рух швидкий – збільшується витримка окремого кадру що сприяє появі ефекту Motion blur. Motion blur означає розмиття зображення при русі [2].

### 1.1.1 Математичне представлення відео обробки

Кожен кадр є зображенням, яке має свою щільність пікселів – відношення сторін кадру. Редагування відео розпочинається із формування кадрів, після чого кожен кадр оброблюється як звичайне 2D зображення. Тобто фільтри та маски накладаються на кожен кадр окремо.

При цифровій обробці кадру (зображення) зазвичай використовується його подання у пам'яті у вигляді матриці пікселів  $f(m_1, m_2)$ . Обробка зображення в загальному випадку полягає у виконанні будь-якого перетворення зазначеної матриці, в результаті якого формується набір її числових характеристик або нове, оброблене зображення –  $g(n_1, n_2)$ . Перетворення може стосуватися значень елементів або їх координат (індексів), виконуватися над матрицею в цілому, групою елементів або над кожним елементом окремо.

Найбільш розповсюдженими фільтрами є зміна яскравості та контрасту. Аргументом функції є колір пікселя первинного зображення. Значення функції є колір пікселя обробленого зображення. Для зміни яскравості або контрасту функція застосовується для кожного пікселя зображення [6-8]. Кожен елемент матриці зображення є 24-бітове значення, де кожен компонент моделі RGB представлений 8-ма бітами. Тому замість інтервалу  $[0, 1]$  використовується інтервал  $[0, 255]$ .

Якщо яскравість і контраст зображення не змінюються в процесі перетворення, то функція має графік, представлений на (рис. 1.5). З рисунка видно, що функція в цьому випадку просто передає на вихід значення свого аргументу.

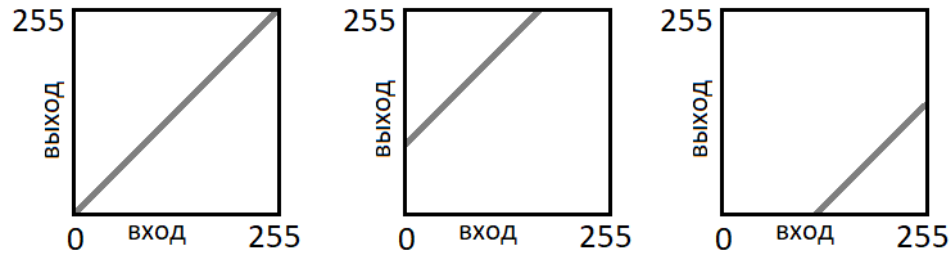


Рисунок 1.5 – Графіки яскравості, збільшення яскравості (посередні) та зменшення яскравості (справа)

Яскравість для даної функції є зсув прямої лінії в вертикальному напрямку. Яскравість зображення збільшується пропорційно зсуву прямої. Якщо пряма зсувається вгору яскравість зображення збільшується, а якщо пряма зсувається вниз – зменшується.

При використанні перетворення контрасту пряма лінія змінює свій нахил. При збільшенні контрасту зображення (рис. 1.6) нахил прямої збільшується, при зменшенні контрасту – зменшується. При цьому зрушення прямої в горизонтальному напрямку означає, що крім контрасту змінюється і яскравість зображення.

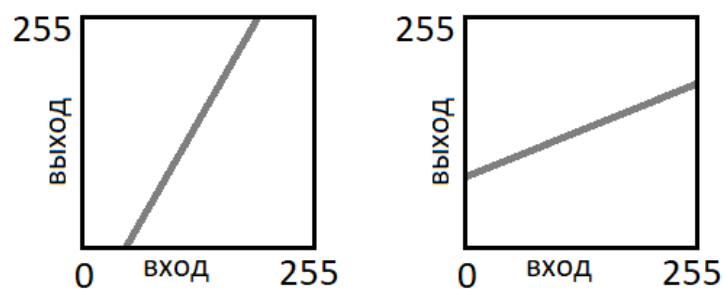


Рисунок 1.6 – Графіки контрасту, збільшення контрасту (зліва) та зменшення контрасту (справа)

Комбінації нахилу і зсуву прямої дозволяють одночасно змінювати і яскравість, і контраст зображення. Перетворення яскравості або контрасту може бути застосовано і до окремих компонентів, наприклад до компоненту

зеленого кольору. Тоді яскравість / контраст будуть змінюватися тільки для червоного компонента, а для інших компонент вони залишаться незмінними. Більш того, можна задавати різні перетворення яскравості або контрасту одночасно для кожного компонента.

В обробці відео також може бути застосовано масштабування. Масштабування зображення дозволяє стиснути або розтягнути його по горизонталі і/або вертикалі. При використанні білінійної інтерполяції колір обчислюється, як зважена сума найближчих чотирьох пікселів вихідного зображення (при збільшенні) або як зважена сума групи пікселів (при зменшенні).

Зазвичай використовується білінійна або бікубічна інтерполяція для зміни масштабу. Бікубічна інтерполяція дозволяє отримати зображення із більш високою якістю, ніж білінійна інтерполяція. Однак слід зауважити, що при подальшому підвищенні порядку інтерполяції якість одержуваного зображення може поліпшуватися незначно. Наведемо найпростішу формулу, яка дозволяє визначити найближчий піксель вихідного зображення (без використання інтерполяції):

$$C_{new}[i][j] = C_{old}[k_1 * i][k_2 * j],$$

де

$$i = \overline{0, H_{old} - 1}, j = \overline{0, W_{old} - 1}, k_1 = \frac{H_{old}}{H_{new}}, k_2 = \frac{W_{old}}{W_{new}}.$$

Параметр  $W$  визначає розмір зображення по горизонталі, вимірюваний в пікселях. Параметр  $H$  визначає розмір по вертикалі. Параметри  $i$  і  $j$  визначають відповідно рядок і стовпець матриці зображення і змінюються в межах висоти і ширини зображення відповідно [9-12].

Перетворення повороту, також як і при розгляді плоских геометричних об'єктів, дозволяє повертати вихідне зображення на заданий кут. Поворот здійснюється навколо центру зображення. При цьому можливі два варіанти повороту, або області зображення, що вийшли за його межі при повороті

відсікаються, а незаповнені частини заповнюються будь-яким кольором. Або розраховується новий розмір зображення на основі кута повороту таким чином, щоб повернене зображення цілком помістилося в нові розміри. Незаповнені частини зображення також заповнюються будь-яким кольором. У будь-якому випадку для розрахунку перетворення повороту може бути використана наступна формула:

$$C_{new}[i][j] = \begin{cases} C_{old}[a][b], a \in [0, H_{old} - 1]^{b \in [0, W_{old} - 1]} \\ C, a \notin [0, H_{old} - 1]^{b \notin [0, W_{old} - 1]} \end{cases},$$

де

$$a = \left\lfloor i * \sin \varphi + \frac{H_{new}}{2} \right\rfloor, \quad b = \left\lfloor j * \cos \varphi + \frac{W_{new}}{2} \right\rfloor,$$

$$i = \overline{0, H_{old} - 1}, \quad j = \overline{0, W_{old} - 1}.$$

У цій формулі параметр  $C$  визначає колір, яким заповнюються порожні ділянки зображення. Параметр  $\varphi$  визначає кут повороту за годинниковою стрілкою в радіанах. Наведена формула округлює перетворені координати. Однак можна використовувати і білінійну інтерполяцію, коли колір пікселя обчислюється як зважена сума квітів чотирьох сусідніх пікселів.

Цифрові фільтри дозволяють накладати на зображення різні ефекти, наприклад: розмиття, різкість, деформацію, шум та інше. Цифровий фільтр є алгоритмом обробки зображення. Велика група цифрових фільтрів має один і той же алгоритм, але ефект, що накладається фільтром на зображення, залежить від коефіцієнтів, використовуваних в алгоритмі.

Згортка є спосіб представлення будь-якого векторного значення скалярним значенням. Існує нескінченна кількість таких способів, багато з яких визначаються деякими коефіцієнтами. Стосовно до обробки зображень векторне значення є колір групи пікселів, а скалярний значення, що отримується на основі згортки, є колір пікселя, одержуваного в результаті застосування до вихідного зображення будь-якого ефекту.

Фільтр характеризується розміром групи пікселів – розміром фільтра і своєї імпульсною характеристикою. Імпульсна характеристика залежить від розміру фільтру і визначається коефіцієнтами фільтра. Також відзначимо, що розмір фільтра завжди є непарним. Коефіцієнти фільтра є деякі скалярні значення, на які множаться значення кольорів пікселів з групи, що відповідає розміру фільтра. Обробка зображення із застосуванням такого роду фільтрів описується наступною формулами:

$$C_{new}[i][j] = \sum_{k=0}^m \sum_{l=0}^n \alpha_{k,l}$$

$$C_{old}[i - m/2][j - n/2].$$

У цій формулі коефіцієнти  $\alpha$  визначають той ефект, який накладає фільтр. Константи  $m$  і  $n$  задають розмір фільтра (він є двовимірним).

Зображення, так само як і звук, може розглядатися як суперпозиція функцій синуса і косинуса з різною амплітудою і фазою. При цьому ці функції є двовимірними, так як саме зображення двовимірне. Високі частоти в зображенні означають різкі зміни яскравості пікселів. Низькі частоти означають плавні зміни яскравості пікселів. Часто виникає завдання збільшення різкості зображення, що означає посилення високих частот. Також може виникати необхідність зменшення різкості і збільшення розмитості, що означає посилення низьких частот. Наступна матриця визначає коефіцієнти цифрового фільтра розміром  $3 \times 3$  пікселі, використовуваного для підвищення різкості зображення.

$$\begin{array}{ccc} -k/8 & -k/8 & -k/8 \\ -k/8 & k+1 & -k/8, \\ -k/8 & -k/8 & -k/8 \end{array}$$

де параметр  $k$  визначає ступінь підвищення контрасту. Зазвичай використовується  $k = 2$ .

Наступна матриця визначає коефіцієнти цифрового фільтра розміром  $3 \times 3$  пікселі, використовуваного для підвищення розмитості зображення.

$$\begin{array}{ccc} 1/9 & 1/9 & 1/9 \\ 1/9 & 1/9 & 1/9. \\ 1/9 & 1/9 & 1/9 \end{array}$$

У даній матриці значення всіх коефіцієнтів однакові. Це практично означає, що кольори групи пікселів  $3 \times 3$  просто усереднюються. Розмір матриці визначає ступінь розмитості. Наприклад, якщо використовується матриця  $5 \times 5$ , де кожен елемент дорівнює  $1/25$ , то розмитість зображення буде більше, ніж при використанні матриці  $3 \times 3$ . Також часто використовується розмиття Гауса [13-16].

З наведених вище прикладів видно, що сума всіх коефіцієнтів дорівнює одиниці. Це основна умова, що використовується при розрахунку фільтрів з кінцевою імпульсною характеристикою. Недотримання цієї умови призводить до того, що контраст зображення після застосування фільтра змінюється.

Для створення переходу між двома кадрами накладається динамічний ефект на перший та на другий кадр. При цьому кадри деформуються за основними правилами обробки зображень. Так звані маски, що накладаються

в реальному часі, демонструють технології додаткової реальності. Ці ефекти є одними із найпопулярніших сучасних методів відео обробки в мобільних додатках [17, 18].

### 1.1.2 Сучасні формати відео

Більшість відеоредакторів підтримує декілька форматів запису. Тобто є можливість обробляти відео записана на різних носіях та обрати формат при збереженні відео після обробки. Досі використовуються такі цифрові формати як MiniDV, HDV та AVCHD.

При MiniDV та HDV запис йде на магнітну стрічку, проте якість запису HDV набагато краще – дозвіл  $1440 \times 1080$  при співвідношенні сторін 16:9. Формат відео AVCHD високої чіткості, назва якого розшифровується як Advanced Video Coding High Definition. Використовується для запису відео у високій чіткості 720p (HD ready), 1080i і 1080p (Full HD) режимах на жорсткі диски або карти пам'яті (SD, microSD і Memory Stick PRO). Якщо запис йде на диски, то це зазвичай перезаписувані AVCHD або Blu-Ray диски.

При завантаженні відео в якусь програму або інтернет ресурс відбувається стиснення. Ступінь стиснення відеопотоку і формат кодека визначається цілями подальшого застосування цього відео. Для високоякісної мовної продукції потрібні відеоматеріали з мінімальними спотвореннями, тобто з малим ступенем стиснення і, відповідно, з великим потоком – для відео стандартної чіткості від 25 Мбіт / с і вище. Для розміщення в інтернеті і для запису на мобільні пристрої застосовуються ефективні кодеки, що дозволяють отримати прийнятну якість зображення з невисоким потоком даних – 1-2 Мбіт/с для відео стандартної чіткості. Відео високої чіткості вимагає великих бітрейтів і, відповідно, підвищених вимог до системних ресурсів комп'ютера або монтажною станцією [19].

MPEG – один з основних стандартів стиснення файлів, розшифровується як Moving Pictures Expert Group. Він має чотири різновиди. MPEG-1 – формат стиснення файлів для CD з посередньою якістю відео (352×240), в даний час практично не використовується. MPEG-2 використовується для цифрового телебачення і DVD-дисків. MPEG-3 (не плутати з технологією стиснення звуку MP3) – в даний час не використовується. Формат MPEG-4 можна отримати за допомогою таких кодеків, як XviD, DivX та ін. Незважаючи на те, що він стискає відео сильніше, ніж MPEG-2, якість картинки цілком пристойне, особливо при використанні кодека H.264.

H.264 – стандарт, що відрізняється високим ступенем стиснення відеоданих при збереженні гідної якості картинки. Чудово підходить для відеороликів, що завантажуються в Мережу.

XviD і DivX – кодекси, що відрізняються високим ступенем стиснення (засновані на стандарті MPEG-4).

Розширення відео (як і будь-якого іншого файлу) представляє собою набір символів, які допомагають операційній системі Windows визначити тип відомостей, що містяться в файлі, а також підібрати підходящу для його відкриття програму[19].

AVI (Audio-Video Interleaved) – дуже популярний формат, точніше кажучи – контейнер. Його розробила компанія Microsoft. У такому контейнері можуть зберігатися дані чотирьох типів – аудіо, відео, текст і midi. У нього можуть входити відео будь-якого формату, починаючи з MPEG-1 і закінчуючи MPEG-4, звуки різних форматів; поєднання кодеків може бути будь-яким. Вміст контейнера AVI можна дізнатися за допомогою різних програм, починаючи з простої VideoToolBox і закінчуючи потужною Adobe Premiere.

WMV (Windows Media Video) – ще один формат від «Майкрософта». Якщо ви створите відеоролик в програмі Movie Maker, яка поставляється разом з ОС Windows, він буде мати саме такий формат.

MOV. Даний формат винайдено компанією Apple Macintosh. Крім відеоданих може містити графіку, анімацію і 3D. Файли цього формату програватимуться в QuickTime Player від все тієї ж Apple.

MKV (Matroska або «Матрьошка») – ще один контейнер, який може містити як відео і аудіо, так і субтитри тощо. MKV володіє відкритим кодом і останнім часом використовується практично повсюдно. Файли даного формату програватимуться на більшості плеєрів, в тому числі на програвачі Windows Media Player, який входить до складу Windows [19].

3GP – формат відео третього покоління для мобільних телефонів. Відеоролики, записані в цьому форматі, відрізняються низькою якістю картинки. Однак малий розмір нівелює цей недолік. Крім того, на маленьких телефонних дисплеях відео в форматі 3gp виглядає більш-менш зрозумілим.

Відеоролики, створювані в Інтернеті, як правило, мають один з наступних форматів: FLV, SWF, RealVideo. FLV (Flash Video) – формат відео, використовуваний цілим рядом місць для розміщення відеофайлів, таких як YouTube, RuTube, Google Video та інші. Анімація або відео в Flash-форматі, створені в програмі Adobe Flash, має розширення SWF (Shockwave Flash). Для програвання таких файлів необхідний браузер з встановленим Flash Player'ом. Ролики в цьому форматі також повсюдно зустрічаються в Інтернеті, де їх можна дивитися прямо з сайту. RealVideo – формат, створений компанією RealNetworks і використовується для телевізійних трансляцій в Мережі. Відеоролики, записані в цьому форматі, мають розширення RM, RA або RAM, відрізняються невеликими розмірами і посередньою якістю [19].

Застосунок TikTok підтримує формати MOV та MP4, відношення сторін кадру 9:16, що в пікселях 1080×1920. Ці правила та вимоги застосовні для медіафайлів, які додаються з Галереї на смартфоні або інших соціальних мереж.

### 1.1.3 Можливості редактора

Редагування відео – програма, що включає в себе набір інструментів, які дозволяють здійснювати нелінійний монтаж відео-та звукових файлів на комп'ютері. Крім того, більшість відеоредакторів дозволяють створювати і накладати титри, здійснювати колірну і тональну корекцію зображення, мікшувати звук і створювати спецефекти. Програми професійного призначення дозволяють синхронізувати звук із зображенням по тимчасовому коду. Зазвичай, у мобільних додатках є можливість захоплення відео замість завантажування готових файлів. З метою економії дискового простору відеопотік при захопленні стискається, тобто кодується із застосуванням алгоритмів компресії.

Функціонал відеоредакторів відрізняється за складністю та набором можливостей. Сучасні мобільні додатки такі як Telegram, Instagram, Facebook впроваджують базові можливості для відео обробки. Зазвичай це завантажування відео, обрізка, зміна стандартних параметрів (яскравість, контраст та інші), доповнення накладеними графічними елементами як додаткового шару, виключення звуку та інше.

Існують спеціалізовані програми для обробки відео на смартфоні, які мають розширений функціонал на рівні професійних застосунків. Розширений асортимент фільтрів та масок, обрізка відео за розміром кадрів та їх кількістю, формування відео із більш ніж одного запису методом склеювання, поворот або віддзеркалювання відео (перетворення застосовується відносно усіх кадрів обраного запису для збереження цілісності відео), розмивання контурів або цілих кадрів. Важливою перевагою сучасних відеоредакторів є можливість хромакею – робити прозорим обраний колір для створення нового шару поверх оригінального ролику, що відтворює ефект додаткової реальності

Окремою перевагою є можливість регулювання звуку. Можна залишити оригінальний звук та накласти додатковим шаром інший, при

цьому змінити гучність кожного. Додаткові шари можна розташувати на будь-якому відрізку стрічки. Аудіо легко синхронізувати за часом із потоком відео. Запис можна змінити за швидкістю відтворення відео. При цьому є шанс погіршити якість відео через велику відстань часу між двома сусідніми кадрами, запис почне смикатися і стане менш плавним та зрозумілим.

Поверх оригінальних кадрів можна розташувати графічні елементи, створенні самостійно або обрані із доступних стікерів, накласти текст. Кожен елемент може бути розташований у будь-якому місці відносно кадру та з'являтися і зникати в обраний час. Графічні додатки можуть бути динамічними, наприклад маска у вигляді падаючих зірок. По факту ця динамічність досягнута шляхом надання можливості додавати не тільки зображення (стікери) а й у форматі GIF (зациклений відеозапис без звукового супроводження). Розмір та нахил кожного елементу можна змінити за бажанням.

Відео стрічка може бути розрізана на декілька окремих стрічок, що дозволяє додати більше переходів між записами. Переходи представляють собою динамічні ефекти стосовно крайніх кадрів двох сусідніх записів. Вони поєднують два записи за допомогою додаткового фільтру під час зміни кадрів. Під час роботи із різними фільтрами є можливість попереднього результату, тобто первинний результат до збереження кінцевого.

Після завершення редагування можна обрати формат та розширення за бажанням, мінімізувати стиснення для збереження якості. Але оригінальний запис як правило залишається найчіткішим із усіх створених роликів після обробки.

## 1.2 Огляд технічних засобів

Обробка відео є розповсюдженим засобом перетворення звичайного відео у справжній фільм із доповненням у вигляді ефектів. Впровадження

функціоналу відео обробки у мобільні застосунки виводять цю сферу на новий рівень. Основними ознаками є зрозумілість, ефективність, легкість в застосуванні та висока якість кінцевого результату.

Розвиток технічних засобів спричинив розвиток мультимедійних застосунків в напрямку оцифрування та розповсюдження відео матеріалів. Багато додатків, таких як соціальні мережі (наприклад Telegram, Instagram, Facebook) впроваджують технології обробки зображень і відео. Більшість користувачів не публікують фото або відео без хоча б якоїсь обробки, а такі напрямки як блогінг та влогінг набирають популярність в наші дні, що спричинило формування попиту в даному напрямку розвитку технологій.

Якщо комп'ютерна обробка відео обростає новим більш складним функціоналом і може бути коректно використана в умілих руках, то відеообробка на смартфоні це більш простий і доступний варіант який може опанувати будь-який користувач.

Відеоредактори для смартфонів відрізняються від комп'ютерних програм тим, що вони дозволяють не тільки завантажувати вже відзняте відео але й захоплювати його в робочому режимі. Також є різниця у форматі запису та його розширенні. Процес обробки відео з технічної сторони та алгоритми такі самі як при комп'ютерній обробці. Мобільні додатки спрямовані на обробку відео мають більш автоматизований функціонал у порівнянні із комп'ютерними програмами. Тобто більшість ефектів, що накладуватимуться на відеоматеріал не вимагають додаткових рухів від користувача та обробляються автоматично за декількома параметрами, такими як яскравість, контраст, додавання шумів, зміна домінантного кольору та інше.

Мобільні додатки із функціоналом редагування відео можна поділити на категорії – спеціалізовані та неспеціалізовані. Різниця полягає в спрямованості додатку. Тобто для спеціалізованих застосунків редагування відео є основним функціоналом. Для неспеціалізованих це лише додаткова функція.

До спеціалізованих додатків можна віднести відеоредактори Inshot, KineMaster та PowerDirector (табл. 1.1). Для порівняння біли обрані застосунки, що є популярними та мають широкий спектр функцій для обробки відео. Їх можна порівняти за набором можливостей із професійними комп'ютерними програмами.

Таблиця 1.1 – Порівняння можливостей спеціалізованих відеоредакторів Inshot, KineMaster та PowerDirector

	Inshot	KineMaster	PowerDirector
Обробка аудіо	-	+	-
Безлімітна кількість стрічок, розрізання та склеювання	+	+	+
Можливість плавного згасання	+	+	-
Screenshot одного кадру запису	-	+	-
Підтримка 4к	+	+	+
Безлімітна кількість переходів	-	+	+-
Створення свого шрифту для текстових вставок	+	-	+
Анімація для додаткового зображення	-	+	-
Діапазон зміни швидкості відео	0,2 – 10	0,2 – 8	0,2 – 8
Хромакей	+	+	+

Inshot позиціонується як додаток для швидкої обробки для соціальних мереж. Він має можливість обрати формат і відношення сторін під конкретний додаток, наприклад Instagram, TikTok або Youtube, та використовувати деякі унікальні для цих додатків фільтри. KineMaster та PowerDirector не мають прив'язки до інших додатків.

За результатом аналізу спеціалізованих на редагуванні мобільних додатків видно, що основний функціонал дуже схожий в усіх обраних популярних відеоредакторах. Деякі застосунки мають ліміти, але при використанні звичайним юзером ця різниця не помітна.

До не спеціалізованих додатків можна віднести ті застосунки, які мають функцію відеоредагування як додаткову, але не основну. Наприклад, соціальні мережі Telegram, Instagram, Snapchat та TikTok (табл. 1.2). Треба пам'ятати, що перш за все ці додатки є соціальними мережами, тому функціонал доволі обмежений. Для більш наочного порівняння додатків цієї групи представлено огляд базових функцій у форматі таблиці.

Таблиця 1.2 – Порівняння можливостей відеоредакторів неспеціалізованих додатків Telegram, Instagram, Snapchat та TikTok

	Telegram	Instagram	Snapchat	TikTok
Обрізка ролику (за таймінгом)	+	-	+	+
Зміна параметрів	+	-	-	-
Додавання масок, графічних елементів та фільтрів	+	+	+	+
Перегляд та застосування фільтрів та масок в реальному часі	-	+	+	+
Додавання кількох фільтрів на один відео запис	+	-	-	+
Зміна відношення сторін кадру	+	-	-	-
Поєднання кількох стрічок	-	-	-	+
Обробка аудіо	-	-	-	+

Telegram є месенджером, що базується на комунікації через повідомлення. При відправці відео його можна відкоригувати за бажанням – обрізати зайвий матеріал, вимкнути аудіо, змінити параметри кадрів

(яскравість, контраст, світло, тіні та інше), додати текст або намалювати будь що поверх відео.

Instagram має додаток для обміну фото та відео матеріалом. Він має схожий функціонал та надає більш широкий асортимент масок та стікерів як додатковий шар поверх відео. Але зображення не можна відригувати за розміром (змінити відношення сторін кадру) та обрізати за часом.

Snapchat дуже схожий на Instagram за змістом, тому що він також представляє собою додаток для обміну фото та відео. Основною фішкою цього додатку є велика кількість масок які застосовуються в реальному часі. Тобто немає потреби робити фото або відео запис поперед завантаженням у Snapchat, він накладає фільтри в реальному часі. Ці фільтри за допомогою розпізнавання образів накладаються на лице як доповнена реальність. Де які маски є навіть анімованими.

TikTok хоча й належить до категорії не спеціалізованих застосунків, але має найширший функціонал серед цієї групи. Його можливості можна порівняти із спеціалізованими редакторами. Це пов'язано із тим що TikTok ґрунтується на передачі тільки відеоматеріалу. Він надає можливість створювати чат або залишати коментарі під записом, але основною функцією є відео редагування та викладення його у мережу.

В результаті аналізу порівняння функціоналу мобільних додатків щодо обробки відео, видно переваги застосунку TikTok перед аналоговими. Деякі функції є унікальними для застосунку, що не повторюється навіть у спеціалізованих додатках.

### 1.3 Постановка задачі дослідження

Розвиток технічних засобів спричинив розвинення мультимедійних застосунків в напрямку оцифрування та розповсюдження відео матеріалів. Багато популярних додатків по типу соц. Мереж (такі як Telegram, Instagram

та TikTok) впроваджують технології обробки зображень і відео. Більшість користувачів не публікують фото і відео без будь-якої обробки, а блогінг та влогінг набирають популярність в наші дні, що спричинило формування попиту в даному напрямку розвитку технологій.

Об'єктом дослідження є мобільний застосунок TikTok, що є соціальною мережею, яка дає можливість знімати музичні відео кліпи, вести прямі ефіри і обмінюватися повідомленнями.

Метою дослідження є виявлення особливостей сучасного функціоналу додатку TikTok та обґрунтування його популярності.

Необхідно вирішити такі задачі:

- провести аналіз існуючих методів обробки відео;
- виявити та порівняти засоби для редагування відео на смартфоні;
- провести аналіз машинного навчання та аналізу даних, застосування технологій в соціальних мережах;
- обґрунтувати сучасність технічної сторони застосунку TikTok;
- створити засобами TikTok демонстраційний відеоролик з використанням алгоритмів машинного навчання.

## 2 ЗАСОБИ МАШИННОГО НАВЧАННЯ

### 2.1 Загальні методи машинного навчання

Сучасні методи машинного навчання полягають у використанні нейронних мереж (НМ) для розподілу даних по параметрах, можливості прогнозувати наступний крок, та розпізнавання образів. Існує багато видів нейронних мереж (додаток А), які класифікуються за призначенням.

На сьогодні відбуваються значні зрушення у сфері машинного навчання, оскільки нейронні мережі, штучний інтелект та інші подібні технології перестали бути тільки лабораторної концепцією. З тих пір як комп'ютери стали досить потужними і доступними, почалася практична робота. І хоча, найбільші комп'ютерно-нейронні мережі здатні повторювати лише 1% діяльності головного мозку, вони вже можуть розпізнавати обличчя на фотографіях, вивчати смаки і звички користувача, і до певного рівня розуміти, що ви говорите [20-22].

Ідеї, відображені в дослідженнях Стефана Гроссберга, покладені в основу багатьох сучасних нейромережових розробок. Вхідна зірка, як показано на рисунку, складається з нейрона, на який подається група входів через синаптичні ваги (рис. 2.1). Нейрон в формі вхідної зірки має безліч входів, яким відповідають ваги і один вихід  $Y$ , що є зваженою сумою входів.

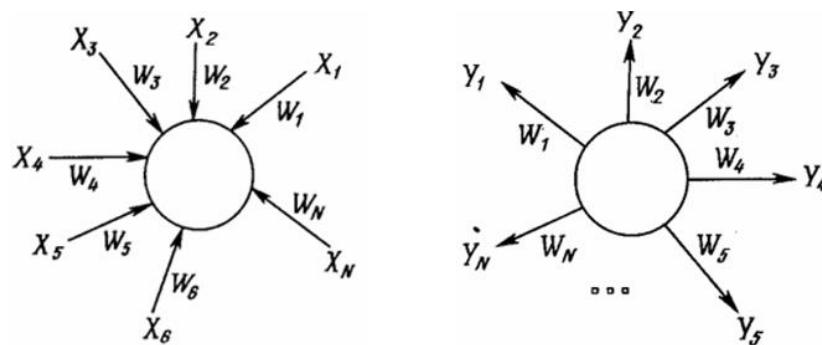


Рисунок 2.1 – Зірки Гроссберга, інтстар (зліва) та аутстар (справа)

Вихідна зірка, праворуч, є нейроном, керуючим групою ваг. Нейрон цього типу має один вхід і безліч виходів з вагами. Вхідні і вихідні зірки можуть бути взаємно з'єднані в мережі будь-якої складності.

Особливістю нейронів в формі зірок Гроссберга є локальність пам'яті. Кожен нейрон в формі вхідний зірки пам'ятає свій, що відноситься до нього і ігнорує інші. Кожній вихідній зірці властива також конкретна командна функція. Образ пам'яті зв'язується з певним нейроном, а не виникає внаслідок взаємодії безлічі нейронів в мережі [23-25].

Вхідна зірка виконує розпізнавання образів, тобто вона навчається реагувати на певний вхідний вектор. Це навчання реалізується шляхом настройки ваг. Вихід зірки визначається як зважена сума її входів. Процес навчання виражається наступним виразом:

$$w_i(t + 1) = w_i(t) + a[x_i - w_i(t)],$$

де  $w_i$  – це вага входу  $x_i$ ;

$x_i$  –  $i$ -й вхід;

$a$  – нормуючий коефіцієнт навчання, який має початкове значення 0,1 і поступово зменшується в процесі навчання.

Добре навчений вхідні зірка буде реагувати не тільки на певний одиничний вектор, але також і на незначні зміни цього вектора. Це досягається поступової налаштуванням нейронних ваг. Відбувається за наступним алгоритмом, для цього ваги налаштовуються відповідно до необхідного цільовим вектором

$$w_i(t + 1) = w_i(t) + b[y_i - w_i(t)],$$

де  $b$  є нормуючий коефіцієнт навчання, який на початку приблизно дорівнює одиниці і поступово зменшується до нуля в процесі навчання.

Нейронна мережа Гроссберга застосовується в мережі зустрічного розповсюдження. Зустрічне розповсюдження має переваги в тих випадках, де довге навчання неможливо. У зустрічному розповсюдженні об'єднані два добре відомих алгоритму: самоорганізована карта Кохонена і зірка Гроссберга.

НМ зустрічного поширення навчається на вибірці пар векторів. Коли мережа навчена, додаток вхідного вектора приводить до необхідного вихідного вектору. Узагальнююча здатність мережі дозволяє відновлювати пари векторів по одній відомій компоненті ( $X$  або  $Y$ ). Це дозволяє використовувати дану мережу для розпізнавання образів, відновлення образів та посилення сигналів. Мережа зустрічного поширення складається з двох шарів нейронів [25-27].

У мережі зустрічного поширення спочатку навчається шар Кохонена. Після визначення переможця його вихідне значення встановлюється рівним 1, а вихідні значення всіх інших нейронів Кохонена приймаються рівними 0. Нейрон-переможець і всі нейрони вихідного шару, з якими він пов'язаний, формують так звану «зірку» Гроссберга (рис. 2.2).

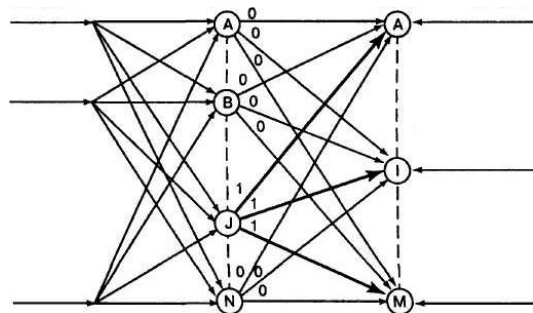


Рисунок 2.2 – Архітектура мережі зустрічного поширення (для спрощення зображення показані не всі зв'язки)

Якщо одного кластеру можуть відповідати різні вихідні нейрони, то для настройки вагових коефіцієнтів «виходить зірки» застосовується ітеративна

процедура навчання з учителем на основі очікуваного вектора  $Y$ . Метою процедури навчання є усереднення очікуваних вихідних значень для елементів кожного кластера. Ваги зв'язків нейрона-переможця з вихідними нейронами налаштовуються згідно з формулою:

$$W^{t+1} = W^t + c(Y - W^t),$$

де  $c$  – малий позитивний коефіцієнт навчання;

$W$  – вектор вагових коефіцієнтів зв'язків нейрона-переможця з вихідним шаром;

$Y$  – очікуваний вихідний вектор.

Зауважимо, що цей алгоритм навчання призводить до посилення зв'язків між вузлом  $J$  шару Кохонена і вузлом вихідного шару тільки в тому випадку, якщо  $i$  – нейрон-переможець, і його вихід дорівнює 1, а бажаний вихід при цьому теж дорівнює 1. Тому даний алгоритм можна вважати різновидом методу навчання Хебба, згідно з яким зв'язок між нейронами посилюється всякий раз при її порушенні.

Самоорганізовані карти Кохонена – це одна із різновидів нейромережових алгоритмів. Основною відмінністю даної технології від інших нейронних мереж є те, що при навчанні використовується метод навчання без учителя, тобто результат навчання залежить тільки від структури вхідних даних. Нейронні мережі даного типу часто застосовуються для вирішення широкого спектру завдань, наприклад відновлення пропусків у даних, аналіз даних або пошук закономірностей [28-31].

Алгоритм функціонування карт із самонавчанням (Self Organizing Maps – SOM) є один з варіантів кластеризації багатовимірних векторів. Прикладом таких алгоритмів може служити алгоритм  $k$ -найближчих середніх ( $k$ -means). Важливою відмінністю алгоритму SOM є те, що в ньому всі нейрони впорядковані в деяку структуру (зазвичай двовимірну сітку). При цьому в ході навчання модифікується не тільки нейрон-переможець, а й його сусіди,

але в меншому ступені. За рахунок цього SOM можна вважати одним з методів проектування багатовимірного простору в простір з більш низькою розмірністю. При використанні цього алгоритму вектора, схожі в вихідному просторі, виявляються поруч і на отриманій карті.

SOM має на увазі використання впорядкованої структури нейронів. Зазвичай використовуються одне і двовимірні сітки. При цьому кожен нейрон є  $n$ -мірний вектор-стовпець:

$$w = [w_1, w_2, \dots, w_n],$$

$$Tw = [w_1, w_2, \dots, w_n]^T,$$

де  $n$  визначається розмірністю вихідного простору (розмірністю вхідних векторів).

Зазвичай нейрони розташовуються у вузлах двовимірної сітки з прямокутними або шестикутними осередками. При цьому нейрони також взаємодіють один з одним. Величина цього взаємодії визначається відстанню між нейронами на карті. На рисунку (рис. 2.3) дано приклад відстані для шестикутної і чотирикутної сіток [31].

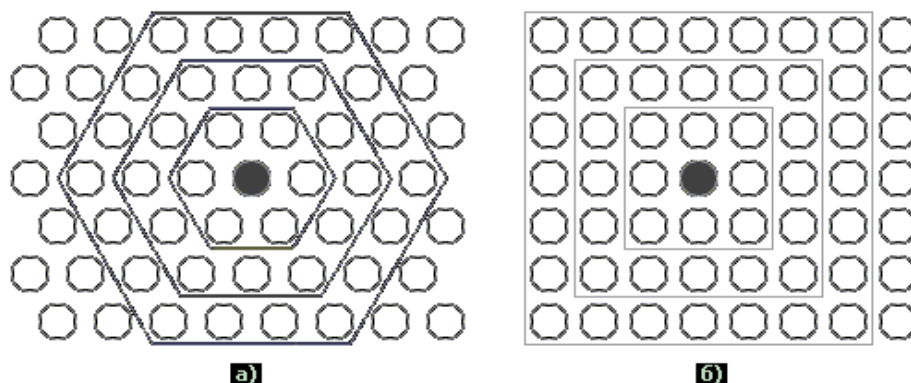


Рисунок 2.3 – Приклад відстані для шестикутної і чотирикутної сіток

При цьому легко помітити, що для шестикутної сітки відстань між нейронами більше збігається з евклідовим відстанню, ніж для чотирикутної

сітки. Кількість нейронів у сітці визначає ступінь деталізації результату роботи алгоритму, і в кінцевому рахунку від цього залежить точність узагальнюючої здатності карти.

При реалізації алгоритму SOM заздалегідь задається конфігурація сітки (прямокутна або шестикутна), а також кількість нейронів в мережі. Деякі джерела рекомендують використовувати максимально можливу кількість нейронів в карті. При цьому початковий радіус навчання в значній мірі впливає на здатність узагальнення за допомогою отриманої карти. У разі, коли кількість вузлів карти перевищує кількість прикладів в навчальній вибірці, то успіх використання алгоритму в великій мірі залежить від відповідного вибору початкового радіуса навчання. Однак, в разі, коли розмір карти становить десятки тисяч нейронів, то час, необхідний на навчання карти зазвичай буває занадто велике для вирішення практичних завдань, таким чином необхідно досягати допустимого компромісу при виборі кількості вузлів [31].

Перед початком навчання карти також необхідно проініціалізувати вагові коефіцієнти нейронів. Вдало обраний спосіб ініціалізації може відчутно прискорити навчання, і привести до отримання більш якісних результатів. Існують три способи ініціювання початкових ваг: ініціалізація випадковими значеннями, коли всім ваг даються малі випадкові величини; ініціалізація прикладами, коли в якості початкових значень задаються значення випадково вибраних прикладів з навчальної вибірки; лінійна ініціалізація. В цьому випадку ваги ініціюються значеннями векторів, лінійно впорядкованих вздовж лінійного підпростору, що проходить між двома головних власними векторами вихідного набору даних. Власні вектора можуть бути знайдені за допомогою процедури Грама-Шмідта.

Навчання складається з послідовності корекцій векторів, що представляють собою нейрони. На кожному кроці навчання з вихідного набору даних випадково вибирається один з векторів, а потім проводиться пошук найбільш схожого на нього вектора коефіцієнтів нейронів. При цьому

вибирається нейрон-переможець, який найбільш схожий на вектор входів. Під схожістю в даній задачі розуміється відстань між векторами, зазвичай обчислюється в евклідовому просторі. Таким чином, якщо позначить нейрон-переможець як  $c$ , то отримуємо формулу.

$$\|x - wc\| = \min_i \|x - w_i\| = \min_i \{\|x - w_i\|\} \{ \|x - w_i\|\}.$$

Після того, як знайдений нейрон-переможець здійснюється коректування ваг нейромережі. При цьому вектор, що описує нейрон-переможець і вектора, що описують його сусідів у сітці, переміщуються в напрямку вхідного вектора. Це проілюстровано на рисунку для двовимірного вектора (рис. 2.4) [31].

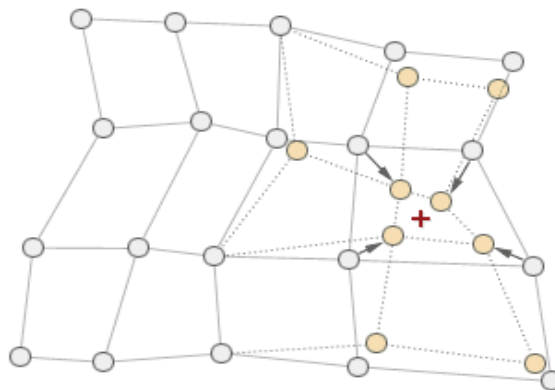


Рисунок 2.4 – Представлення двовимірного вектора

Координати вхідного вектора відзначені хрестом, координати вузлів карти після модифікації відображені сірим кольором. Вид сітки після модифікації відображений штриховими лініями. При цьому для модифікації вагових коефіцієнтів використовується формула:

$$w_i(t + 1) = w_i(t) + hci(t) * [x(t) - w(t)],$$

де  $t$  позначає номер епохи (дискретний час). При цьому вектор  $x(t)$  вибирається випадково з навчальної вибірки на ітерації  $t$ . Функція  $h(t)$  називається функцією сусідства нейронів. Ця функція є незростаюча функцією від часу і відстані між нейроном-переможцем і сусідніми нейронами в сітці. Ця функція розбивається на дві частини: функцію відстані і функції швидкості навчання від часу, де також використовується положення нейрона в сітці [26-31].

Навчання складається з двох основних фаз: на початковому етапі вибирається досить велике значення швидкості навчання і радіуса навчання, що дозволяє розташувати вектора нейронів відповідно до розподілу прикладів у вибірці, а потім проводиться точне підстроювання ваг, коли значення параметрів швидкості навчання багато менше початкових. У разі використання лінійної ініціалізації первісний етап грубого підстроювання може бути пропущений [31-33].

Так як алгоритм SOM поєднує в собі два основних напрямки – векторне квантування і проектування, то можна визначити і основні застосування цього алгоритму. Дану методику можна використовувати для пошуку та аналізу закономірностей у вихідних даних. І після того, як нейрони розміщені на карті, отримана карта може бути відображена. Розглянемо різні способи відображення отриманої карти.

При цьому методі відтворення отриману карту можна представити у вигляді листкового пирога. Кожен шар якого є розмальовку, породжену однієї з компонент вихідних даних. Отриманий набір розмальовок може використовуватися для аналізу закономірностей, наявних між компонентами набору даних. Після формування карти ми отримуємо набір вузлів, який можна відобразити у вигляді двовимірної картини. При цьому кожному вузлу карти можна поставити у відповідність ділянку на рисунку, чотирьох або шестикутний, координати якого визначаються координатами відповідного вузла в решітці (рис. 2.5).

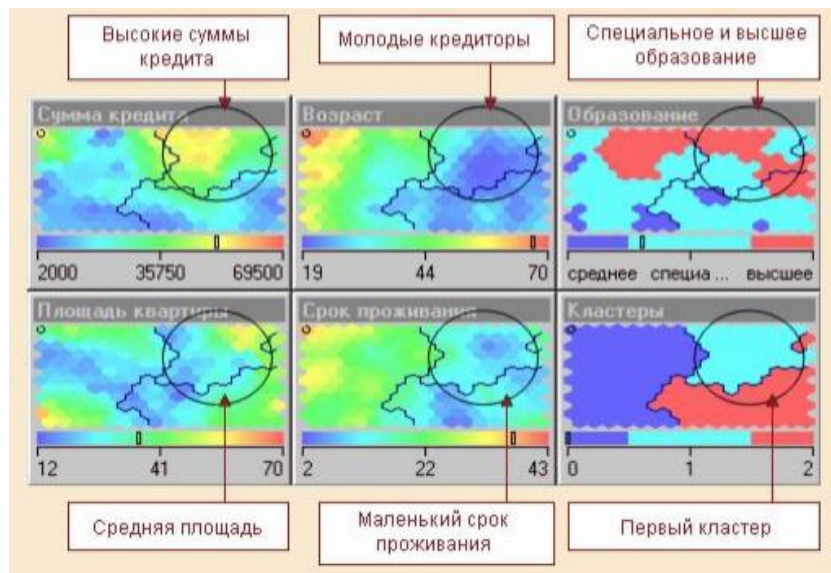


Рисунок 2.5 – Розмальовка, породжена окремими компонентами

Отримані розмальовки в сукупності утворюють атлас, що відображає розташування компонент, зв'язку між ними, а також відносне розташування різних значень компонент [31-34].

Кластером буде група векторів, відстань між якими всередині цієї групи менше, ніж відстань до сусідніх груп. Структура кластерів при використанні алгоритму SOM може бути відображена шляхом візуалізації відстані між опорними векторами (ваговими коефіцієнтами нейронів). При використанні цього методу найчастіше використовується уніфікована матриця відстаней (*u*-matrix). При використанні цього методу обчислюється відстань між вектором ваг нейрона в сітці і його найближчими сусідами. Потім ці значення використовуються для визначення кольору, яким цей вузол буде відмалювали. Як результат, бачимо декілька груп – кластерів, сформованих за алгоритмом SOM (рис. 2.6).

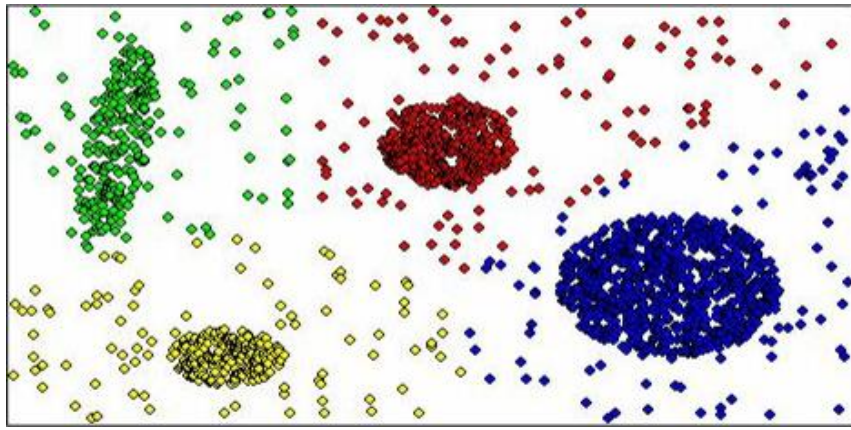


Рисунок 2.6 – Відображення кластерів

Нейронна мережа Хопфілда (Hopfield network, HN) – це повнозв’язна нейронна мережа із симетричною матрицею зв’язків. Під час отримання вхідних даних кожен вузол є входом, в процесі навчання він стає прихованим, а потім стає виходом. Мережа навчається так: значення нейронів встановлюються відповідно до бажаного шаблону, після чого обчислюються ваги, які в подальшому не змінюються. Після того, як мережа навчилася на одному або декількох шаблонах, вона завжди буде зводитися до одного з них (але не завжди – до бажаного). У кожного нейрона є свій поріг активації, що залежить від значення, при проходженні якого нейрон приймає одне з двох значень (зазвичай – 0 або 1). Така мережа часто називається мережею з асоціативною пам’яттю. Тобто, ця мережа, отримуючи не повну групу вхідних даних, може відновити їх до повної [31-35].

Машина Больцмана (Boltzmann machine, BM) дуже схожа на мережу Хопфілда, але в ній деякі нейрони позначені як вхідні, а деякі – як приховані. Машина Больцмана – це стохастична мережа, да навчання проходить за методом зворотного поширення помилки або за алгоритмом порівняльної розбіжності. Загалом процес навчання дуже схожий як у мережі Хопфілда.

Обмежена машина Больцмана (restricted Boltzmann machine, RBM) виходячи з назви дуже схожа на машину Больцмана, отже й на мережу Хопфілда. Нейрони цієї мережі одного типу не пов’язані між собою. Обмежену машину Больцмана можна навчати як FFNN, але замість прямої

передачі даних і зворотного поширення помилки потрібно передавати дані в двох напрямках: спершу в прямому, потім в зворотному. Після чого проходить навчання за методом прямого і зворотного поширення помилки [35-37].

Ланцюги Маркова (Markov chains, MC або discrete time Markov Chains, DTMC) – це попередники машин Больцмана (BM) і мереж Хопфілда (HN). Зміст пояснюється наступним чином: прораховується шанс потрапити в один з наступних вузлів, при знаходженні в розглянутому вузлі. Кожний наступний стан залежить тільки від попереднього. Хоча ланцюги Маркова не є нейронною пережею, але вони дуже схожі. Також ланцюги Маркова не обов'язково повнозв'язні [37-39].

Автокодувальник (autoencoder, AE) не зовсім є фундаментально іншим алгоритмом та схожий на FFNN, тому що нагадує трохи інший спосіб використання FFNN. Основною ідеєю є автоматичного кодування інформації, тобто стиснення даних але не шифрування. Мережа навчається методом зворотного поширення помилки. Подаючи вхідні дані і задаючи помилку дорівнює різниці між входом і виходом [37-39].

Розріджений автокодувальник (sparse autoencoder, SAE) – виглядає як протилежний простому кодувальнику. При аналізі даних ця мережа використовує незвично велику кількість вузлів що призводить до роздування мережі у центрі. Подібний тип мереж добре працює в задачах де необхідно опрацювати дані із великою кількістю дрібних параметрів. Тому цю мережу не можна навчати як звичайний автокодувальник. Окрім вхідних даних необхідно подати ще і спеціальний фільтр розрідженості, який пропускає тільки певні помилки [35-39].

Варіаційні автокодувальники (variational autoencoder, VAE) схожі за архітектурою на AE, але використовуються трохи іншим способом. Тіку мережу навчають наближенню імовірнісного розподілу вхідних зразків, що нагадує роботу машин Больцмана. Проте, варіаційні автокодувальники спираються на алгоритми Байєсова, бо мова йде про імовірнісні висновки і

незалежності, які інтуїтивно зрозумілі, але складні в реалізації. Ця мережа враховує незв'язність подій в різних місцях.

Шумозаглушуючі автокодувальники (denoising autoencoder, DAE) – це АЕ, в які вхідні дані подаються в зашумленому стані. Помилка обчислюється як для звичайних автокодувальників, і вихідні дані порівнюються з зашумленими. Це дозволяє мережі вчитися таким чином, щоб звертати увагу на широкі та вагомні ознаки замість маленьких, оскільки вони можуть бути незначними та змінюватися разом з шумом [35-39].

Згорткові нейронні мережі (convolutional neural networks, CNN) і глибинні згоровані нейронні мережі (deep convolutional neural networks, DCNN) сильно відрізняються від інших видів мереж. Зазвичай вони застосовуються для класифікації об'єктів на зображеннях, в рекомендаційних системах та при обробці мови. Саме ці алгоритми реалізовані в більшості систем сучасних додатків.

Такі мережі зазвичай використовують «сканер», що не парсує усі дані за один раз. Наприклад, якщо є зображення  $200 \times 200$ , не буде відразу оброблятися всі 40 тисяч пікселів. Замість це мережа вважає квадрат розміру  $20 \times 20$  (зазвичай з лівого верхнього кута), потім зрушиться на 1 піксель і вважає новий квадрат, і т.д. Ці вхідні дані потім передаються через згоровані шари, в яких не всі вузли з'єднані між собою. Ці шари мають властивість стискуватися з глибиною, причому часто використовуються ступеня двійки: 32, 16, 8, 4, 2, 1. На практиці до кінця CNN прикріплюють FFNN для подальшої обробки даних. Такі мережі називаються глибинними (DCNN) [39-41].

## 2.2 Використання технологій у сучасних застосунках

Більшість алгоритмічних прийомів, що використовуються в підходах глибокого навчання, відомі і доступні вже деякий час. Значна частина

останніх інновацій у цій галузі стосується того, щоб зробити ці методи доцільними для реальних застосувань. Це передбачає розробку та впровадження архітектур, які можуть виконувати ці методи, використовуючи розумну кількість ресурсів за розумний проміжок часу.

Компанія Netflix запустила проект рекомендаційної системи на основі концепції під назвою «глибоке навчання». Це цілий набір методів, досліджень і практик в області штучного інтелекту, головним завданням якого все так же є поки ще нездійсненне бажання людей змусити комп'ютер мислити, як людський мозок.

Перший рівень обробки здійснюється через створення окремих моделей для різних країн, що пов'язано із місцевим менталітетом. Однак, підготовка одного екземпляра насправді передбачає навчання та тестування декількох моделей, кожна з яких відповідає різним комбінаціям гіперпараметрів. Це є другий рівень, на якому процес може бути розподілений. Цей рівень особливо цікавий, якщо існує багато параметрів для оптимізації, і у вас є хороша стратегія їх оптимізації, як байєсівська оптимізація за Гауссовими процесами. Єдиним зв'язком між прогонами є налаштування гіперпараметрів та метрики оцінки тесту.

На відміну від Google і Facebook, які експериментують зі штучним інтелектом на власних потужностях, Netflix планує запускати свої алгоритми на серверах Amazon. Це показує, що мережеві хмари дозволяють середньому і великому бізнесу конкурувати з корпораціями на рівних. Впровадження крайових рішень, таких як використання графічних процесорів для навчання великомасштабних нейронних мереж, може бути непростою справою. Залучення загальнодоступної хмари AWS може мати очевидні переваги за умови догляду за налаштуваннями та використанням ресурсів екземпляра.

Також Netflix за рахунок нейронних мереж вже експериментує з сюжетами, які залежать від рішень і дій глядача. Нові можливості ще не доступні, але незабаром заміна акторів стане реальним.

Оскільки Instagram є частиною компанії Facebook, вони розглядатимуться разом як від одного розробника. Це обґрунтовується також тим, що в цих застосунках використовуються ідентичні або схожі технології нейронних мереж.

Багато AI-інструментів виконуються на смартфонах, однак це може привести до скорочення терміну служби батареї пристрою і зниження його продуктивності. Тому рішення про те, чи повинна нейронна мережа запускатися на пристрої або в хмарі, важливо не тільки для кінцевих користувачів, але і для підприємств, що розробляють AI-інструменти. Технологія AutoScale автоматизує прийняття рішень про розгортання нейронних мереж [42].

При кожному запуску нейромережі AutoScale враховує архітектурні характеристики алгоритму і можливі відхилення в часі його виконання, такі як рівень сигналу Wi-Fi, Bluetooth і LTE, динамічне масштабування частоти процесора і об'єм використаної пам'яті. Потім він вибирає апаратні засоби – процесори і відеокарти, – які максимізують ефективність енергоспоживання при максимальній якості виконання алгоритму. Потім AutoScale запускає нейросеть, стежачи за показниками пристрою, і визначає, наскільки вибір обладнання поліпшив ефективність.

Щоб перевірити AutoScale, автори розробки провели експерименти на трьох смартфонах, вимірявши енергоспоживання на кожному з них: Xiaomi Mi 8 Pro, Samsung Galaxy S10e і Motorola Moto X Force. Для імітації виконання алгоритму в хмарі вони підключили смартфони до сервера через Wi-Fi. Запуск нейромережі на пристрої вони імітували за допомогою планшета Samsung Galaxy Tab S6, підключеного до смартфонів через Wi-Fi Direct. Для кожного тесту дослідники розробили три сценарії: перевірка комп'ютерного зору без потокової передачі, коли нейросеть аналізувала фото з камери; перевірка комп'ютерного зору на відео в реальному часі з камер; переклад тексту, введеного з клавіатури.

Розробник повідомляє, що у всіх сценаріях AutoScale перевершує базові показники, зберігаючи при цьому низьку затримку виконання алгоритму і високу продуктивність. Крім того, використання AutoScale підвищило енергоефективність в 1,6-9,8 разів для різних сценаріїв. Більш того, для AutoScale треба було тільки 0,4 МБ пам'яті, що становить 0,01% від обсягу ОЗУ (це приблизно 3 ГБ) типового смартфона середнього класу.

Facebook як один зі світових лідерів в розробці технологій штучного інтелекту знайшов прикладне застосування для цих програм: фільтрація спілкування людей. Передові інновації впроваджують в соціальній мережі Instagram [42].

Instagram розробляє і представить нові функції, які «спрямовані на стримування онлайн-булінг» (цькування). Це відбувається в той час, коли в усьому світі зростають заклики посилити контроль за такими платформами, як Facebook і Twitter, на тлі критики за поширення ненависті та фейкових новин, а також за широко поширену агресію і грубість в коментарях. Компанія заявила, що протягом багатьох років використовує AI для моніторингу цькування і шкідливого контенту, а тепер представляє інтерфейс безпосередньої взаємодії цього AI з користувачами.

Ця технологія миттєвого виявлення, яка спрацьовує коли користувач намагається опублікувати щось образливе. Система буде видавати попередження перед публікацією. Інша нова функція – «Обмеження» (Restrict) – спрямована на обмеження образливих коментарів у стрічці користувача.

Створення штучного інтелекту для боротьби з грубістю – це означає навчання машин розуміння проблеми зі складними нюансами. Інженери Instagram вперше використовували DeepText в 2016 році для пошуку спаму. У 2017 році его навчили знаходити й блокувати образливі коментарі, включаючи расові образи. До середини 2018 року програму – по суті, набір нейромереж, – навчили знаходити в коментарях також й цькування, тобто

умисну грубість. У жовтні 2018 року компанія оголосила, що система AI буде аналізувати не тільки коментарі, але й самі пости [42].

Сьогодні Instagram використовує для сканування контенту три окремих класифікатори цювання. Один навчений на розпізнавання ображаючого тексту, другий – фотографій, третій – відео. Класифікатори вже існують в нормальному режимі: працюють, шукають і позначають контент. Проте, вони поки знаходяться «на досить ранній стадії розробки». Тобто допускають багато хибнопозитивних і помилково негативні спрацьовувань.

Успіх клікбейтних заголовків Toutiao і вірусних рекомендацій TikTok забезпечують потужні алгоритми машинного навчання. Ці алгоритми – не публікуються бо є секретним ключем успіху компанії. Аналізуючи поведінку користувача, вони швидко підлаштовуються під його смаки і пропонують саме той контент, який його втягне. Чим більше користувачів, тим більше даних для навчання – і тим ефективніше працюють рекомендації.

TikTok використовує механіку трендів подібну Twitter. Щоб привернути увагу, потрібно опублікувати відео ролик, який містить згадку трендового в даний момент хеш-тега або ключового слова [43].

Terark, що інтегрував TikTok у 2019 році – це технологія, яка дозволяє швидко знаходити дані у величезних стислих масивах даних. У Тік-Ток додається і стискається (розмір файлу) мільйони відео щогодини. Terark вирішує задачу швидкого складання індивідуально плей-лист користувача, вибираючи з сотень мільйонів відео.

Супереч Instagram та YouTube, для TikTok не потрібно жодного перегляду чи підписників, щоб перетворити звичайне відео на «вірусне». У ньому використовується інформаційна стрічка, що працює на основі штучного інтелекту, яка вчиться інтересам користувача з часом користування. Тобто, чим більше її використовують, тим більше вона дізнається. Для прикладу, якщо з'являється новий користувач у програмі, і TikTok показав йому відео, і відвідувач не переглядає ціле відео, то після отримаєте менше рекомендацій щодо таких типів відео і навпаки.

По мірі прогресу в епоху цифрових технологій здібності до видобутку інформації та персональних даних швидко розвиваються. В мобільних додатках всі дані збираються із дозволу користувача, що надає згоду на обробку персональних даних при завантаженні або реєстрації [43].

ТікТок потребує доступ до персональних даних так само як і інші застосунки по типу соціальних мереж. А саме, доступ до ваших контактів, щоб програма могла запропонувати зареєстрованих користувачів, яких ви можете знати. Snapchat, Instagram, Facebook, LinkedIn тощо, також все це роблять, коли людина підписується на їхні послуги. Це найпростіший спосіб знайти людей у мережі, яких відвідувач вже знає.

Одним із важливих параметрів персоналізації є надання даних про місцезнаходження. Це необхідно для того щоб надавати актуальні новини для користувача. Наприклад, якщо людина знаходиться в Україні, програма буде надавати новини переважно пов'язані із цією країною. Twitter, Snapchat, Instagram та інші додатки надають пріоритет своїй місцю знаходження користувача для підбору контенту сторінки «популярне».

Також при аналізі даних враховується вік користувача. Згідно із Законом про захист дітей в Інтернеті (CIPA), дітям, які не досягли 13-річного віку, заборонено мати облікові записи в Інтернеті, оскільки в США заборонено збирати дані про дітей. Ось чому кожен додаток для соціальних мереж, включно з ТікТок, не дозволяє реєстрацію нікому до 13 років.

Для обробки зібраних даних використовуються технології глибокого навчання, що дозволяють аналізувати складні дані, такі як зображення, відео, аудіо, мова. Основна статистика дозволяє робити регресію між двома змінними, але більш досконалі моделі дозволяють співвідносити кілька визначених входів і мати один вихід. За допомогою машинного навчання програми здатні робити це у великих масштабах, використовуючи безліч входів та мільйони рядків даних. Це проста концепція, але широко застосовується, наприклад такі алгоритми як випадкові ліси, підтримуючі векторні машини та багат шарові перцептрони (рис. 2.7).

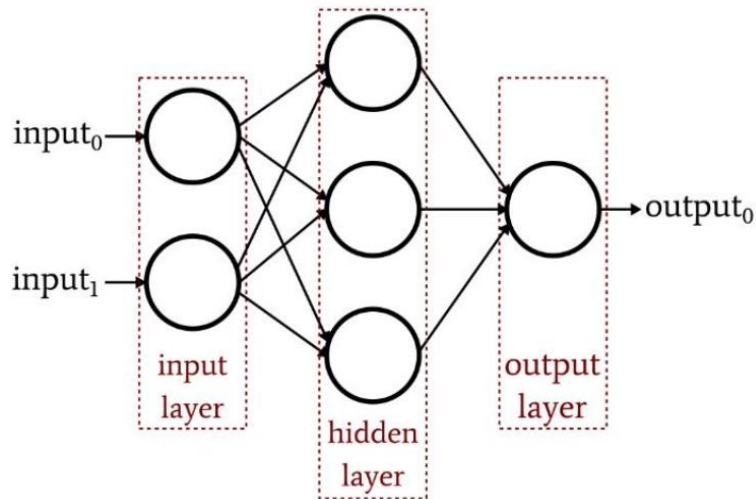


Рисунок 2.7 – Приклад багатошарового перцептрону

Останній перцептрон зазвичай називають штучною нейронною мережею, що робить класифікацію даних. Ця логіка полягає в глибокому аналізі отриманих даних та прогнозуванні. Даний алгоритм є стандартним та загальним для більшості мереж із машинним навчанням, та відрізняється у мобільних додатках «начинкою» у прихованих шарах із різними реалізованими методами аналізу даних. Відстежити ефективність алгоритму можна за допомогою статистичних показників [43].

За статистикою близько 90% усіх світових даних було здобуто за останні два роки. Це пов'язано із поширенням соціальних мереж не тільки як месенджерів, але й платформ для передачі медійних даних у вигляді зображень та відео. Логічно, що для автоматизування обробки великої кількості даних задіюються алгоритми глибокого навчання [44].

Подібно до того, як Netflix та Instagram використовують взаємодію з медіа як способи оцінити інтерес, це робить і TikTok. Однак те, що відрізняє TikTok, – це його здатність вимірювати нетривіальну інформацію. Тобто додаток може аналізувати музику у фоновому режимі, підписи та сам вміст контенту.

### 3 ВИВЧЕННЯ ТА ЗАСТОСУВАННЯ ФУНКЦІОНАЛУ ЗАСТОСУНКУ ТІКТОК

#### 3.1 Огляд можливостей застосунку

Ще кілька років тому вважалося, що ніхто не зможе створювати більш ефективні рекомендаційні алгоритми, ніж Google і Facebook з мільярдами користувачів. Зараз пости в TikTok набирають в чотири рази більше лайків, ніж пости в Instagram з тією ж кількістю переглядів.

Існує 2 версії TikTok: Китайська і Міжнародна. Зрозуміло, що перша доступна тільки в Китаї, друга – в усьому іншому світі. Це пов'язано з юридичними обмеженнями Китаю. Для аналізу додатку в роботі використовувалась глобальна версія.

Контент в TikTok завжди займає весь екран – це новий простір цифрового залучення, укупі з нескінченною стрічкою. Додаток, який фокусує увагу користувача на єдиному ролику на весь екран – отримує підвищену тривалість перегляду, ніж те, де відео займає лише частину екрану і конкурує за увагу з іншим контентом.

Розвиток технологій смартфонів зробило широкодоступними фото або відео зйомку. А редактори реальності, такі як фільтри і цифрові маски дозволяють генерувати нескінченні варіанти контенту.

##### 3.1.1 Обробка відео

Застосунок TikTok дозволяє використовувати готові відеоматеріали або самостійно відзняти відео через застосунок. Під час захоплення відео є можливість накласти маску як додаткову реальність. Також можна зробити паузу, змінити маску після чого продовжити запис. В системі відео із новою

маскою буде зберігатися як окреме від першого. Тобто його можна буде обрізати за тривалістю та змінити порядок записів відокремлених паузою.

Після утворення цілісної стрічки відео, є можливість накласти додаткові статичні фільтри, які накладуватимуться поверх усього ролику, та динамічні фільтри які можна розташувати на стрічці на визначений час. Статичні фільтри відрізняються від динамічних тим що представляють собою набір обробки кадрів за контрастом, яскравістю, зміною кольору та інше. Динамічні фільтри це різного роду викривлення переходу між кадрами або маска із додатковою реальністю (зірки, сніжинки, сердечка та інше).

Окремо від фільтрів можливо накласти додаткові елементи – стікери, смайли або текст. До кожного елемента можна налаштувати час його появи у кадрі. Всі елементи розташовуються в центрі та їх можна змістити, повернути та змінити розмір за потреби. Текст також налаштовується за шрифтом, кольором, вирівнюванням відносно осей  $x$  та  $y$ .

Звісно можна залишити рідне звукове супроводження, але є можливість записати окремо аудіо і накласти його на відео запис. Після цього «рідне» аудіо можна вимкнути або налаштувати гучність обох аудіо записів. Також ТікТок розширив функціональність за допомогою заохочення використання фонові музики в відео. В наявності є велика бібліотека з музикою та популярними голосовими записами. Тобто є можливість накласти на відео не тільки музичне супроводження, а й голос іншої популярної людини чи інший аудіо запис, який не є музичним виконанням.

При завантаженні готового відеоматеріалу є функція синхронізації яка обрізає та склеює кадри таким чином щоб музика співпадала із зміною відео матеріалу. При цьому кожен відео запис обрізається чітко за часом таким чином, що тривалість повного відео залежить не тільки від кількості відзнятого матеріалу, а й від обраної аудіо композиції. При склеюванні кількох стрічок не має можливості налаштувати перехід між ними, але це можна зробити за допомогою динамічних фільтрів.

Для опробування функції синхронізації музики із відео в ході практичного застосування функцій відеообробки були використані готові відзняті відео та фотоматеріали університету ХНУРЕ (додаток Б). Музика була обрана випадковим чином, за рекомендацією додатку TikTok як одне із популярних звукових супроводжень (рис. 3.1). Музика автоматично синхронізується із обраними записами, які були обрізані за часом таким чином, щоб співпадати із ритмом обраної композиції.

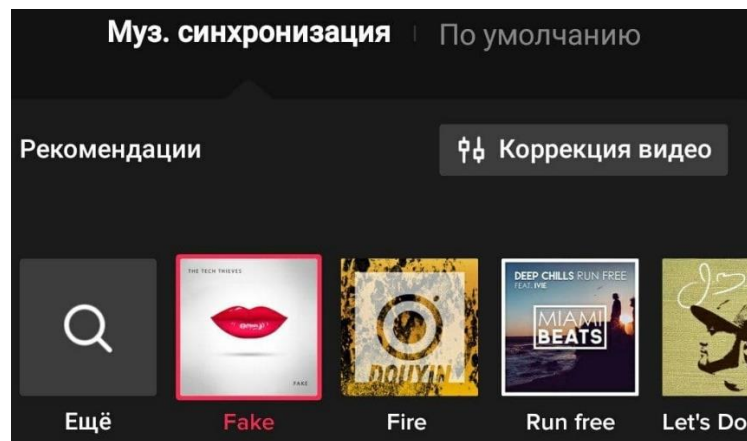


Рисунок 3.1 – Функція музичної синхронізації та вибірка рекомендованих композицій

Для структурної та смислової цілісності відео матеріал було розташовано в конкретному порядку у режимі корекції відео, таким чином час на кожен ролик було встановлено автоматично. Тобто, якщо спочатку на ролик А, який розташований на першому місці було відведено 2,1 секунди, а на ролик Б, що другий за порядком – 2,8 секунд, після зміни їх розташування на загальній стрічці час буде змінено на відповідний (рис. 3.2).

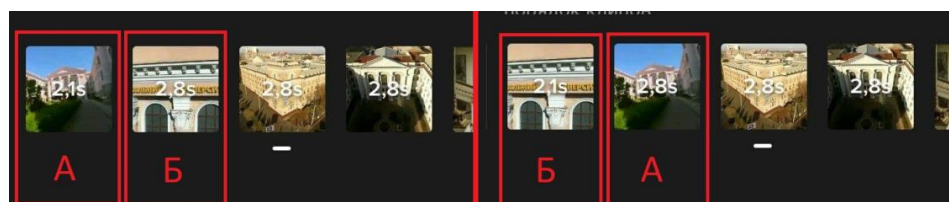


Рисунок 3.2 – Зміна часу роликів залежно від розташування у стрічці

Якщо ролик Б виявиться коротшим за 2,1 секунди, алгоритм музичної синхронізації візьме тривалість першого ролику за основу та змінить час на всі інші записи у стрічці таким чином, щоб збереглися ритм та загальна тривалість відео. Якщо ролик довший за виділений для нього час то можливо обрати який фрагмент буде показано, а які частини запису будуть обрізані (рис. 3.3). Під час налаштування при необхідності можна перегорнути відео запис.

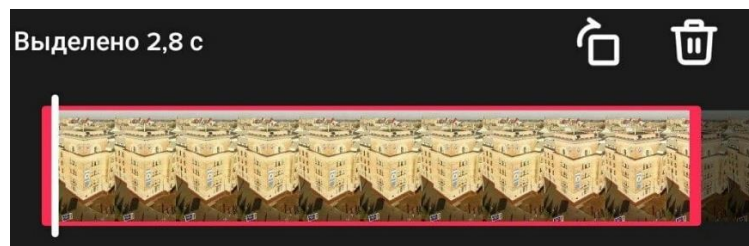


Рисунок 3.3 – Вибір частини запису для показу одного ролику

Для статичних кадрів (фото) час показу встановлюється залежно лише від ритму музики. Анімація фото представляється у вигляді віддалення уявленої камери від центру фото, тобто зміна масштабу від зменшеного до повного розміру зображення.

Під час обробки окремих записів можна додавати чи видаляти відео або фото матеріали зі стрічки, що вплине на загальний час відео. Після налаштування окремих фрагментів та музики йде етап обробки цілої стрічки відео (рис. 3.4). На основній сторінці редактору є не активні функції обрізки та гучності через використану функцію синхронізації музики, яка налаштовує ці параметри автоматично. Функція музики теж не активна, та помічена галочкою як вже використана.

Налаштування фільтрів передбачає обробку усіх кадрів методами обробки зображення (зміни контрасту, яскравості, домінуючого кольору та інше). Для використання фільтрів необхідно обрати один зі списку

запропонованих та налаштувати інтенсивність застосування обраного фільтру на шкалі від 0 до 100 (рис. 3.5).

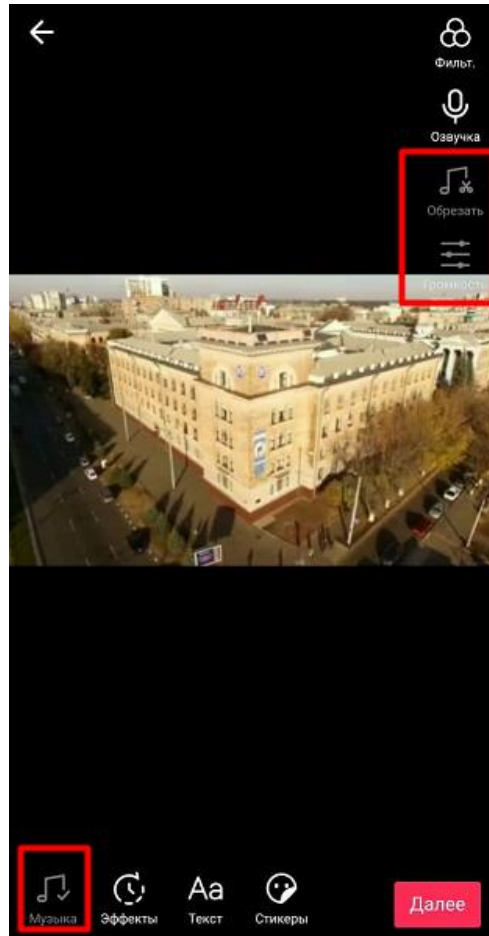


Рисунок 3.4 – Основний редактор відео (активні та не активні елементи)

Навіть після обраного музичного супроводження можливо додати власний звук. Запис голосу буде відразу показаний на стрічці червоним кольором (рис. 3.5). Ця функція не використовувалася у практичній роботі.

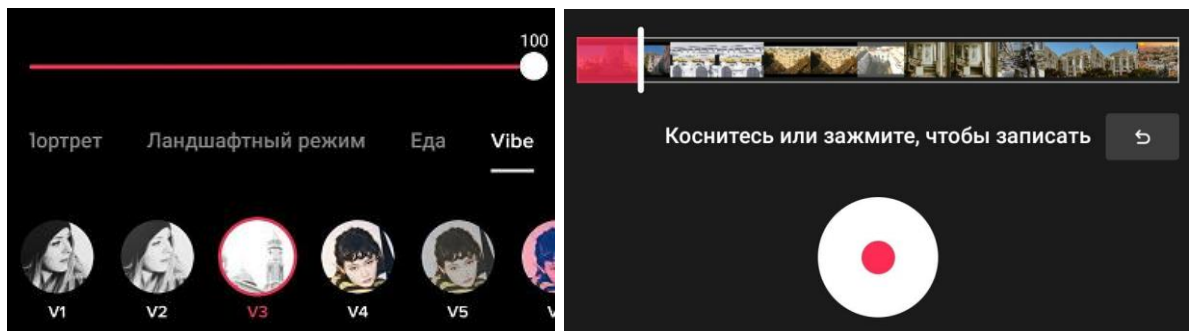


Рисунок 3.5 – Функція фільтрів (зліва) та запису звуку (справа)

Динамічні ефекти поділяються на групи: ті які мають власну тривалість (Transition) та ті які можна використовувати без обмежень часу (Visual, Sticker, Split, Time). Під час використання ці фільтри позначаються на стрічці різними кольорами та не переключаються між групами. Наприклад, ефекти з групи Visual не відображаються під час роботи із ефектами Transition (рис. 3.6).

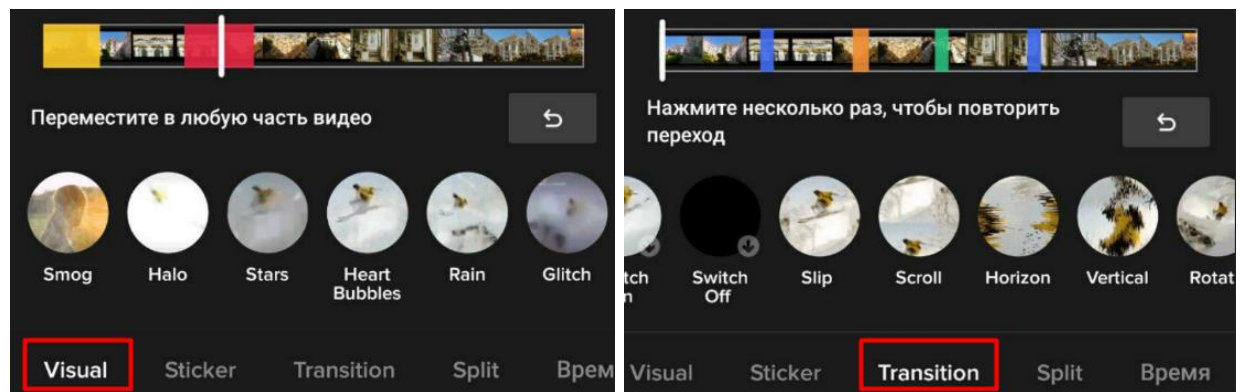


Рисунок 3.6 – Динамічні ефекти Visual (зліва) та Transition (справа)

В роботі було також використано функцію додавання тексту. Редактор тексту дозволяє обрати колір, шрифт, рамку або виділення тексту та вирівнювання відносно горизонталі (рис. 3.7). Розташування та розмір напису налаштовується окремо на записі.

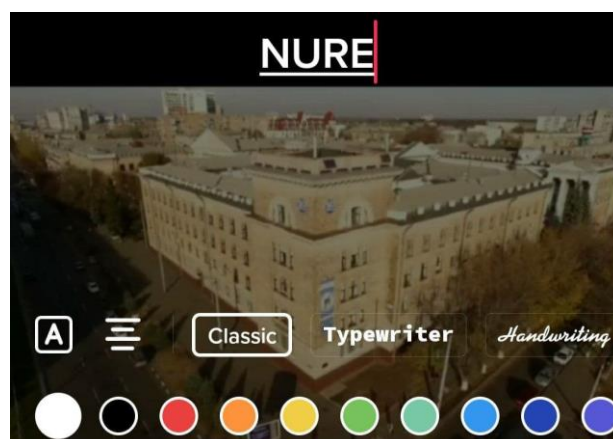


Рисунок 3.7 – Редактор тексту

Після закінчення обробки відео та додавання усіх необхідних елементів, настає завершальний етап. Необхідно придумати назву відео, додати характеризуючі хештеги за потреби просування ролику, обрати відповідні параметри конфіденціальності та обрати обкладинку – стопкадр, що буде видно в режимі перегляду записів. Якщо увімкнути параметр завантаження відео на смартфон, воно відразу збережеться у пам'яті телефону та відео матиме логотип TikTok і аккаунт людини, що створила ролик.

### 3.1.2 Застосування штучного інтелекту

Існує не так багато прозорості щодо алгоритмів, але все ж існує кілька теорій, заснованих на досвіді користувачів та аналізі експертів. Відповідно до теорії партій, коли відео завантажується, воно демонструється групі партій з різними ідеологіями або простими словами на основі кількох відмінних факторів, таких як історія перегляду, місце розташування, особисті уподобання. Доля вашого контенту базується на цих різних партіях; Повторний перегляд, вподобання, спільний доступ та коментарі – це ті невеликі вхідні дані, які використовує алгоритм для ескалації балів вашого вмісту.

Якщо контент користувача подобається групам, тоді його розширюють для більшої аудиторії того самого процесу мислення, що робить його «вірусним». Як допоміжний фактор, чим більше користувач бере участь у тенденціях, тим більше це його просуває.

У світлі принципу рейтингу влади, доля викласти «вірусне» відео залежить від перших кількох записів. Якщо ті набирають певних поглядів і вподобань, то це створює основу для творця, і його єдиною роботою буде якомога більше взаємодіяти з аудиторією. Отримавши рейтинг авторитетів,

користувач опиниться в центрі уваги і буде на кілька рівнів вище початківців, хоча все ще новачком.

Дотримуючись принципу імпульсу затримки, деякі користувачі радили не видаляти будь-яку частину вмісту, незалежно від того, наскільки вона погана. Якщо вміст не приносить ніякої користі, і TikTokери починають бути менш активними на платформі, програма намагатиметься вдосконалити їхній вміст самостійно через деякий час, стимулюючи їх створювати більше для своєї аудиторії. Вдосконалення може бути як збільшення випадків попадання до рекомендованих.

Немає підтвердженого робочого процесу для алгоритму TikTok, але на основі даних, розміщених компанією можна зробити деякі висновки. Як тільки відео завантажується на TikTok, високопродуктивний алгоритм штучного інтелекту використовує NLP та комп'ютерний зір для аналізу відео. Він проаналізує кожен частину викладеного відео, включаючи аудіо, субтитри та метадані (опис, хештеги), щоб створити розуміння змісту та контексту відео.

Штучний інтелект керує всіма платформами вмісту Bytedance. Розробники створили інтелектуальні машини, здатні розуміти та аналізувати текст, зображення та відео, використовуючи технологію обробки природною мовою та технологію комп'ютерного зору. Це дозволяє програмі обслуговувати користувачів вмістом, який їм здається найбільш цікавим, і надавати можливість творцям ділитися важливими моментами у повсякденному житті з глобальною аудиторією.

За допомогою комп'ютерного зору, програма виявляє об'єкти розташовані у кадрі. Технологія NLP також оцінює аудіо для більшої точності сприйняття контенту. Опис відео, хештеги та навіть назва музичного супроводження аналізуються у сукупності.

Після перевірки відео за правилами та положеннями, вміст буде передано невеликій частині аудиторії, і відбудеться оцінка на основі того, як вибіркова група користувачів відреагувала на цей вміст. Кожен

відстежуваний показник має можливий пов'язаний бал. Але повторна публікація відео не зробить його «вірусним», оскільки цей алгоритм може розрізнити зайві записи. Як тільки загальний бал для вмісту буде гідним, він підштовхне відео до ширшої аудиторії (рис. 3.8).

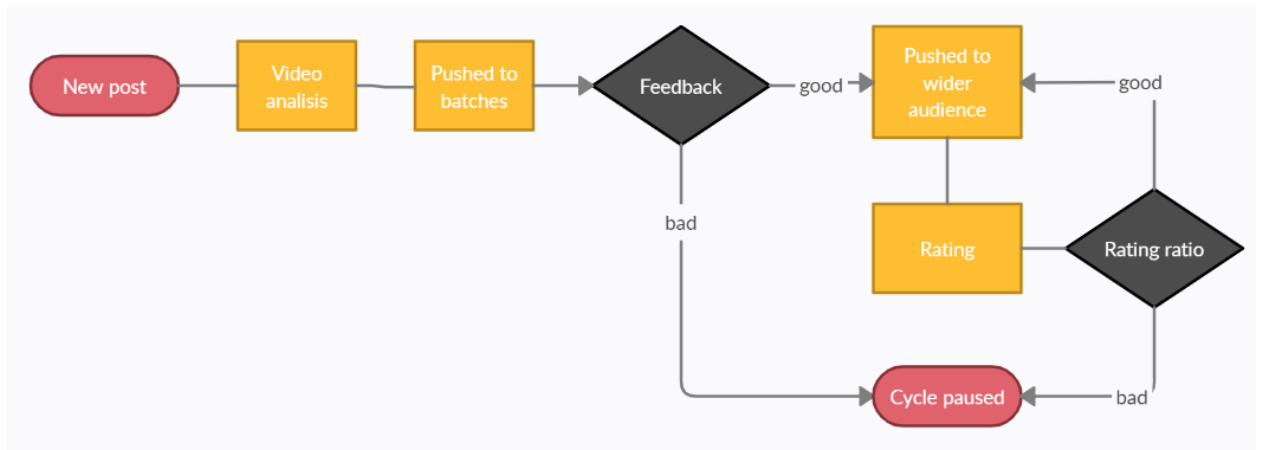


Рисунок 3.8 – Схема обробки інформації нового творця

Якщо вміст не стає «вірусним», цикл не закінчується, а просто призупиняється, і відповідно до теорії затримки імпульсу, він може повернутися назад, коли користувач почне залишати коротший слід у програмі. Це ланцюгова реакція, яка буде тривати і надалі, залучаючи більше творців контенту і, отже, більше аудиторії. В цьому суть цієї платформи, яка може затягувати своїх користувачів протягом тривалого часу.

ТікТок знижує бар'єр для входу, не примушуючи своїх користувачів реєструватися. Додаток починає дізнаватися про відвідувача якомога більше. Перша послідовність відео, які він споживає, допомагає алгоритмам визначити смак людини. Перш за все, щоб втримати нового відвідувача у програмі якомога довше, додаток покаже лише відео, які подобаються ширшому колу аудиторії, таким чином намагаючись зберегти низький рівень виходу.

Після того, як користувач потрапляє в цю екосистему, алгоритми відстежуватимуть кожну метрику та створюватимуть оцінку профілю для кожного жанру. Кожного разу, коли він переглядає відео або споживає всю

тривалість відео, алгоритм бере до уваги ще кілька відео подібного типу для вашого профілю (рис. 3.9).

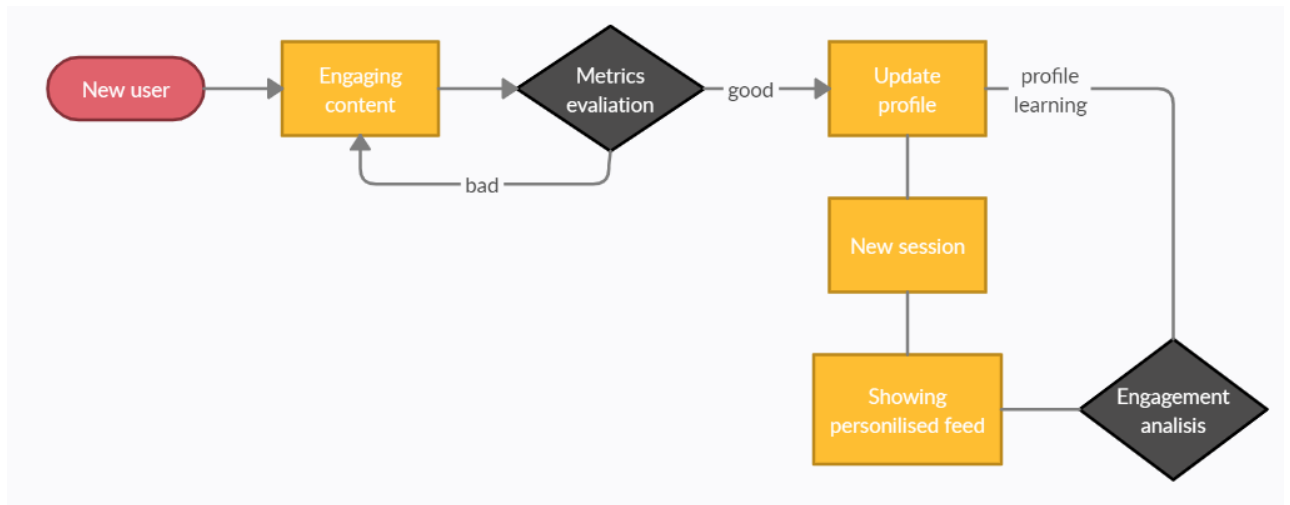


Рисунок 3.9 – Схема обробки інформації нового користувача

Разом з даними метрик, алгоритм також збиратиме особисту інформацію, таку як місцезнаходження, історія пошуку Інтернету, вік, щоб створити більш точний персоналізований канал, тим самим утримуючи користувача в екосистемі якомога довше. Унікальність такого алгоритму полягає в тому, що він може залучити користувачів у всьому світі з різних вікових груп завдяки своїй винятковій системі рекомендацій щодо вмісту [43].

На TikTok мільйони вмісту, що завантажуються користувачами щодня. Практично неможливо відстежувати кожен вміст вручну, тому перевірка проводиться машиною, і якщо вміст позначений нечутливим до протоколів компанії, він передається для перевірки людиною.

Після проходження людського огляду вміст перевіряється на дублювання, далі передається невеликій партії користувачів для початкового зворотного зв'язку. На основі системи метрик надається оцінка вмісту, що визначає долю контенту. Якщо оцінка перевищує деякий поріг, тоді відео розповсюджується для ширшої аудиторії, але заспокоюється приблизно через тиждень, щоб користувачі постійно відчували новий контент.

Але якщо оцінка зворотного зв'язку низька, цикл тут не зупиняється, а навпаки, він призупиняється, і додаток спостерігає за поведінкою користувача та його активним часом, якщо програма відчуває високу швидкість виходу, тоді цикл знову стане активним і збільшить вміст тим самим заохочуючи творця створювати більше і більше залучати на платформі [43].

Цей алгоритм діє таким чином, що забезпечує новим та цікавим персоналізованим контентом кожного користувача. Він спостерігає за відвідувачем та покращуватимете рекомендаційну стрічку відповідно до зафіксованих інтересів людини, що не дає нудьгувати та заохочує більшу аудиторію.

В роботі було застосовано алгоритм штучного інтелекту під час редагування відео – автоматична синхронізація музики та відео записів. TikTok не має повних композицій, лише певні частини запису. Це пов'язано із лімітом на тривалість відео ролику, який не може перевищувати одну хвилину. Тому щоб не займати пам'ять великою кількістю аудіо записів вони були скорочені від 10 секунд до 60 секунд.

Кожен аудіо запис із листу доступних у додатку має тактовий набір маркерів розташованих за часом на стрічці музики. Ці маркери означають певні тактові зміни в композиції. При використанні функції синхронізації музики із відео маркери розташовані в аудіо записі прив'язуються до кожного початку наступного відео.

Рекомендаційний лист аудіо записів підбирається для кожного користувача індивідуально за великою кількістю параметрів. Враховуються фактори місцезнаходження, статі, віку, переглянутого контенту та поміток «Сподобалось», та популярністю використання конкретної композиції. Цей лист постійно оновлюється, з кожним початком редагування він буде вже інакшим ніж в попередній час. Це пов'язано із динамікою використання додатку.

### 3.2 Порівняння функціоналу з іншими застосунками

TikTok належить компанії ByteDance, Китай. Сьогодні цей додаток один з найдорожчих стартапів в світі з оцінкою в \$ 75 мільярдів та у 2018 року став найдорожчим стартапом за всю історію, таким чином обійшовши сервіс таксі Uber який був найдорожчим стартапом до появи TikTok.

TikTok популярний і в країнах СНД. 8 мільйонів користувачів з різних країн користуються додатком хоча б раз на місяць, 3 мільйони роблять це щодня. При цьому майже половина користувачів завантажує якийсь контент – це дуже високий показник.

Прийнято вважати, що Китай копіює західні технології. Але в останні роки ситуація змінилася. Тепер глобальні компанії копіюють успішні китайські продукти. Феноменальна популярність TikTok – лише частина глобальної тенденції. Для аналізу ефективності роботи машинного навчання у додатку було переглянуто статистичні дані.

Зараз ByteDance витрачає \$ 3 мільйони в день на просування TikTok на американському ринку. Завдання просте – за всяку ціну відібрати аудиторію у американських соцмереж. В TikTok немає програми виплат для авторів, як на YouTube, а рекламні інструменти обмежені. Відсутність агресивної монетизації пояснюється цілком раціонально: розробники платформи не хочуть псувати атмосферу креативності. Така тактика працює – на даний момент TikTok практично монополізував ринок вертикальних розважальних відео і став головною альтернативою стрімко старіючим американським соцмережах [45].

Нещодавно TikTok купив британський стартап Jukedeck, що виробляє машинну музику. Це позбавить компанію від проблем з правовласниками, які підняли ціни на свій контент для TikTok в рази. У розпорядженні компанії – не тільки потужні алгоритми рекомендацій і мільярд користувачів, які безкоштовно виробляють контент, але і нескінченна бібліотека машинної музики – у будь-якому жанрі і на будь-який смак

Instagram має функцію історій – «stories», які зараз важливіші за пости, тому що несуть повсякденну актуальну на цей час інформацію та дійсні тільки 24 години. Ці stories схожі за контентом на відео з TikTok, але після завершення терміну придатності цих історій вони завершають набирати перегляди в той час коли відео у TikTok продовжують бути актуальними більш ніж одну добу. Це та розширений функціонал допомагають набирати більшу популярність, що приводить до збільшення користувачів.

Важливим фактором у здатності TikTok підтримувати авторів буде стійкість додатку до автоматизації. У 2020 алгоритми TikTok ефективно визначають перезаливи і не виводять їх в рекомендації.

Instagram обмежує використання музики в Історіях. TikTok – заохочує, тому надав користувачам вбудований в редактор доступ до величезної бібліотеки аудіозаписів. І запровадив унікальну механіку – запис свого відео на звук іншого автора.

За статистикою глобальної версії застосунку TikTok аудиторія представлена наступним чином (рис. 3.10): 26% Сполучені Штати, 6,3% Росія, 3,5% Індія, 3,2% Великобританія. По світу бачимо позитивний приріст. В якості пріоритетних ринків були обрані ті, в яких популярний Instagram – Росія, США, Великобританія та Індія [45].

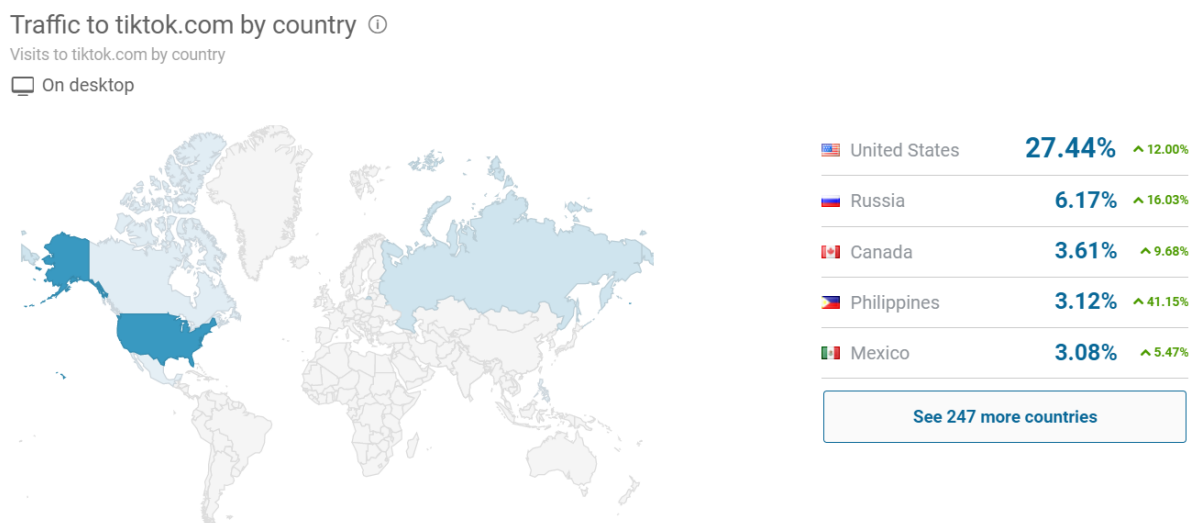


Рисунок 3.10 – Кількість користувачів у світі та приріст

За офіційною статистикою в середньому приріст користувачів за останні 6 місяців більше ніж 10 відсотків та загальна кількість відвідування перевищує 500 мільйонів (рис 3.11).

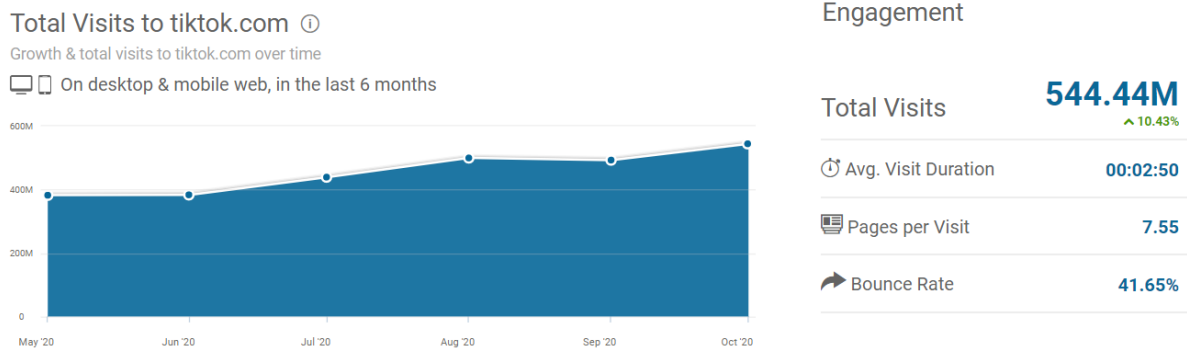


Рисунок 3.11 – Глобальний трафік застосунку за останні 6 місяців

В середньому за відвідування переглядається більше 7 відео та сеанс триває 2,5 хвилини. В середньому одне відео триває 30 секунд. За статистикою, користувач додатка витрачає 52 хвилини в день, з яких 46 хвилин витрачається на перегляд роликів, та останні 6 хвилин на спілкування і коментування. Це означає що додаток отримує інформацію для аналізу про майже 100 відеороликів, що було переглянуто користувачем. Також виділяються основні категорії інтересів користувачів ТікТок – комп'ютерні технології, ігри, електроніка та програмування, соціальні мережі, онлайн товариства.

Було проаналізовано трафік використання застосунку та посилання на відео з ТікТок. За оцінкою ресурсів що посилаються на контент з ТікТок за останній місяць (рис. 3.12) більшість трафіку приходять з месенджерів та пошукових систем, а найчастішою цільовою аудиторією є користувачі таких додатків як Facebook, Google, YouTube, WhatsApp.

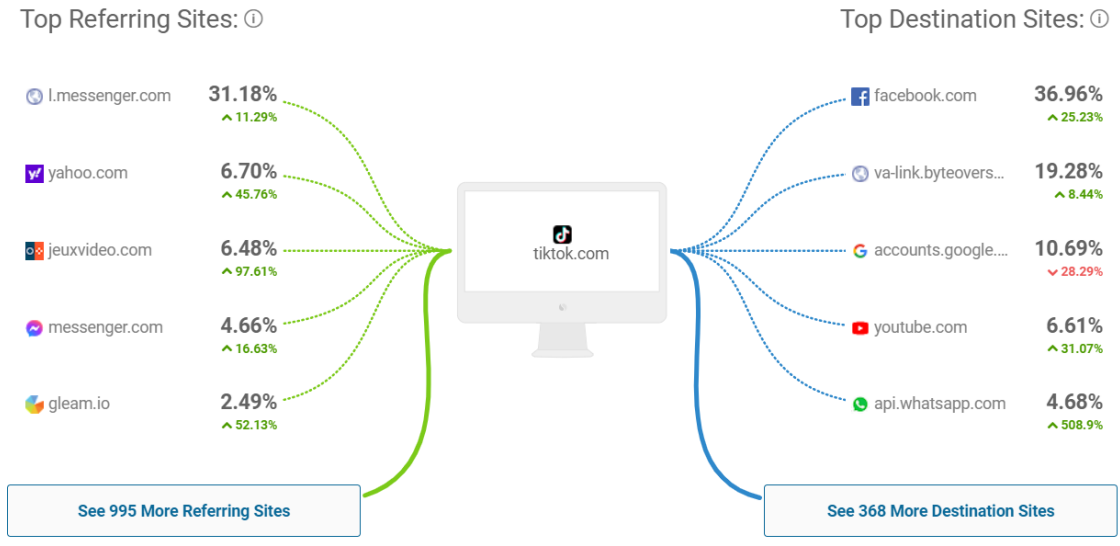


Рисунок 3.12 – Частка популярних ресурсів з посиланням та цільових додатків за останній місяць (жовтень)

За популярністю користувачів інших додатків соціальних мереж виділяються з найпоширенішим додаванням трафіку (рис. 3.13) Facebook, YouTube, Вконтакте, Twitter та WhatsApp. Ці соціальні мережі використовуються найчастіше для посилання на контент з ТікТок за останній місяць (жовтень на момент аналізу) [45].

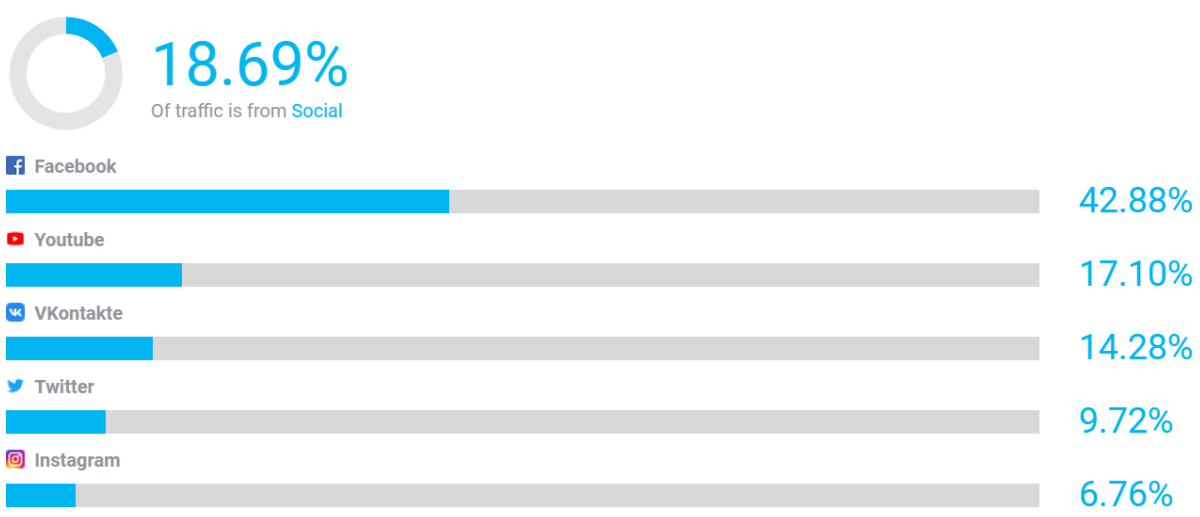


Рисунок 3.13 – Трафік з соціальних мереж за останній місяць (жовтень)

Сьогодні зростання TikTok не має ознак зупинки. Остання статистика TikTok показує, що станом на квітень 2020 року популярний відеододаток було завантажено понад два мільярди разів у всьому світі як в Apple App Store, так і в Google Play. Іншими словами, TikTok зміг подвоїти кількість завантажень лише за рік – яскравий ознака стрімкої популярності програми (рис. 3.14) [46].

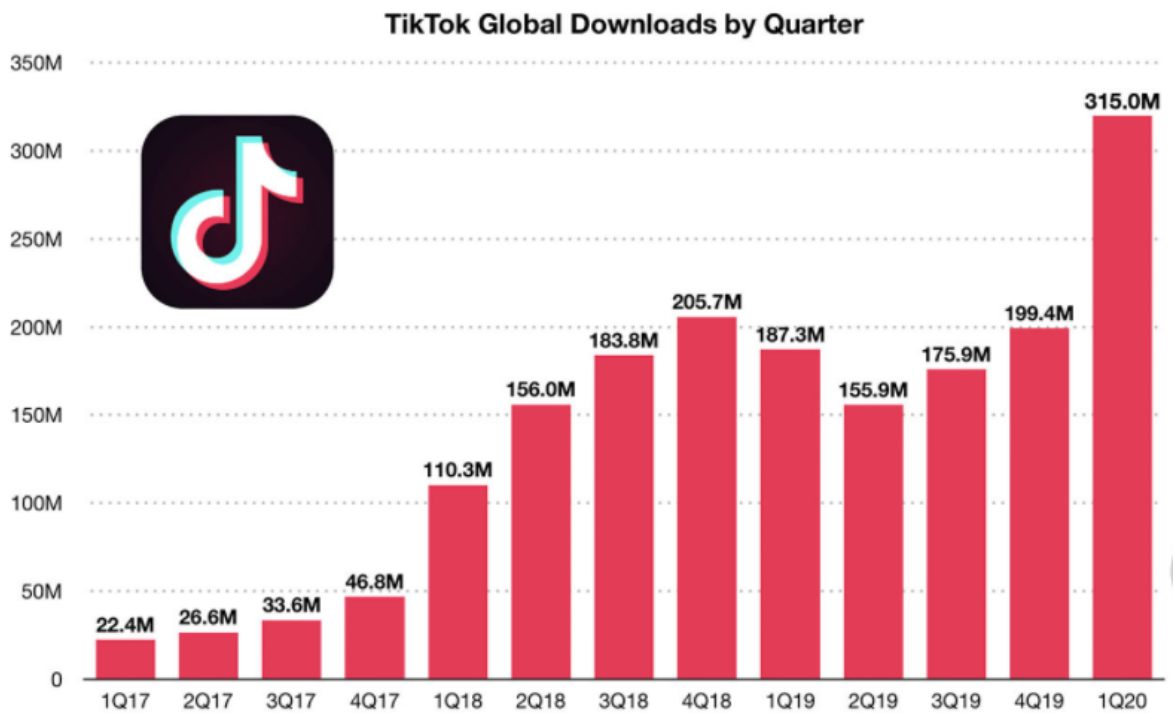


Рисунок 3.14 – Завантаження додатку по кварталам з 2017 року

Остання віха настає лише через п'ять місяців після того, як TikTok перевищив 1,5 мільярда завантажень. У першому кварталі 2020 року він згенерував найбільше завантажень будь-якої програми за весь час (рис. 3.15), накопичивши понад 315 мільйонів завантажень через App Store і Google Play.

Для порівняння із іншими схожими платформами мас-медіа, TikTok займає першу позицію, а Facebook та Instagram ділять між собою другу та третю позиції. Також видно стабільну популярність додатку Snapchat, що займає п'ятий рядок в рейтингу [46].

## Top Social Media Apps Worldwide for January 2020 by Downloads



Overall Downloads	App Store Downloads	Google Play Downloads
1  TikTok	1  TikTok	1  TikTok
2  Facebook	2  Instagram	2  Facebook
3  Instagram	3  Facebook	3  Instagram
4  Likee	4  Kuaishou	4  Likee
5  Snapchat	5  Snapchat	5  Snapchat
6  Helo	6  Twitter	6  Helo
7  Twitter	7  Wesee	7  VMate
8  Kuaishou	8  Pinterest	8  Twitter
9  Pinterest	9  Vigo Video	9  Pinterest
10  VMate	10  WeChat	10  Kuaishou

Note: Does not include downloads from third-party Android stores in China or other regions.

Рисунок 3.15 – Кількість завантаження додатків у січні 2020 року

Основна мета будь-якої платформи соціальних медіа - залучити своїх користувачів, і що дивно, ТікТок може залучати своїх середніх нових користувачів до 10 хвилин, що втричі перевищує можливості Instagram.

Це знижує бар'єри для входу та навіть не вимагає від вас реєстрації, після встановлення програми вона починає відстежувати вас з цього моменту та починає рекомендувати вам відео на основі цих факторів.

Майже 90 відсотків усіх користувачів ТікТок отримують доступ до програми щодня. Мало того, вони надзвичайно активні в додатку. Дослідження, яке спостерігало поведінку користувачів ТікТок протягом одного місяця, показує, що 68 відсотків користувачів ТікТок переглядають чуже відео, а 55 відсотків завантажують власні відео [46]. Також існує різниця у звичках перегляду та обміну даними, причому перші мають чітку перевагу серед користувачів ТікТок. Деякі користувачі можуть віддати перевагу ТікТок як джерелу розваг, але, незважаючи на те, що відвідувачі активніше переглядають наявний вміст, все одно є вражаюча кількість людей, які діляться власними відео. Це також демонструє творчу свободу та простоту ТікТок як соціальної платформи.

## ВИСНОВКИ

У рамках атестаційної роботи було проаналізовано методи обробки відео, виявлено недоліки та переваги сучасних відео редакторів та складність їх функціоналу. Після проведеного огляду сучасних додатків із функціоналом обробки відео можна визначити лідера серед групи не спеціалізованих на відео редагуванні, додатків – застосунок TikTok. Можливості цього застосунку можна порівняти із професійними редакторами.

Сучасні методи машинного навчання полягають у використанні нейронних мереж для розподілу даних по параметрах, можливості прогнозувати наступний крок. Не зважаючи на те, що більшість методів машинного навчання вже відома де-який час, вони постійно застосовуються в сучасних додатках для підбору реклами та рекомендацій. Ефективність впровадження цих технологій залежить від якості проведеного аналізу над метою та прорахунку комбінацій вже відомих методів.

TikTok опереджає інші мобільні додатки у машинному навчанні, тому що не вимагає реєстрації для підбору рекомендацій. Він починає працювати з завантаження та відстежувати інтереси користувача з першого запуску.

В ході роботи створено відео ролик засобами TikTok, використано широкий спектр функцій включаючи алгоритми машинного навчання. Під час аналізу розповсюдження додатку та його популярності було виявлено переваги застосунку серед інших соціальних мереж і великий потенціал як нового стартапного проекту.

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Латкин, А. В. (2016). ХАРАКТЕРИСТИКА МУЛЬТИМЕДИЙНОЙ РЕКЛАМЫ. Информационные технологии в прикладных исследованиях: сб. мат. и, 71.
2. Зиновьев, И. В., & Альберта, Г. ВЫРАЗИТЕЛЬНЫЕ СРЕДСТВА ЭКРАНА КАК ЭЛЕМЕНТЫ МАНИПУЛЯТИВНЫХ ТЕХНОЛОГИЙ (НА ПРИМЕРЕ НОВОСТНЫХ ПРОГРАММ ФЕДЕРАЛЬНЫХ ТЕЛЕКАНАЛОВ).
3. Поляков, В. А., & Романов, А. А. (2017). Разработка и технологии производства рекламного продукта.
4. Что такое двухпотокое цифровое редактирование и монтаж видео? // Компания «СТОИК». URL: [https://www.ixbt.com/divideo/dstream\\_ded.html](https://www.ixbt.com/divideo/dstream_ded.html).
5. Частота кадров (Frames per Second) FPS для видеонаблюдения. // Stronghold. URL: <https://stronghold.com.ua/blog/chastota-kadrov-frames-per-second-fps-dlja-videonabljudeniya-opisanie-i-vybor-optimalnogo-znachenija/#>.
6. Путятін, Є. П., Гороховатський, В. О., & Матат, О. О. (2006). Методи та алгоритми комп'ютерного зору: навч. посіб. Харків: ТОВ «Компанія СМІТ».
7. Гороховатский, В. А. (2014). Структурный анализ и интеллектуальная обработка данных в компьютерном зрении.
8. Gorokhovatskyi, V., Gorokhovatskyi, O., Putyatin Ye. & Peredrii O. (2018, August). Quantization of the Space of Structural Image Features as a Way to Increase Recognition Performance. In 2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP) (pp. 464-467). IEEE.
9. Gorokhovatskyi, V.O., Tvoroshenko, I.S., and Peredrii O.O. (2020) Image classification method modification based on model of logic processing of bit description weights vector, Telecommunications and Radio Engineering, 79(1), pp. 59-69.

10. Гороховатський В.О., Гадецька С.В. (2020) Статистичне оброблення та аналіз даних у структурних методах класифікації зображень (монографія), Харків, ФОП Панов А.Н., 128 с., DOI: 10.30837/978-617-7859-69-6.

11. Гороховатский, В. А., Пуятин, Е. П., & Столяров, В. С. (2017). Исследование результативности структурных методов классификации изображений с применением кластерной модели данных. *Радіоелектроніка, інформатика, управління*, (3), 78-85.

12. Gorokhovatskiy, V. A., & Putyatin, Y. P. (2009). Image Likelihood Measures on the Basis of the Set of Conformities. *Telecommunications and Radio Engineering*, 68(9).

13. Гороховатський, В. А. (2008). Ієрархія просторових відносин структурних ознак в задачах зіставлення візуальних об'єктів. Системи управління, навігації та зв'язку.-К .: Центральний наук.-досл. інститут навігації і управління, 85-89.

14. Gorokhovatskiy, V. A., & Zamula, A. A. (2016). Employment of Intelligent Technologies in Multiparametric Control Systems. *Telecommunications and Radio Engineering*, 75(19).

15. Гороховатский, В. А. (2003). Распознавание изображений в условиях неполной информации. Харьков: ХНУРЭ.

16. Gorokhovatskiy, V. A., & Peredriy, Y. O. (2008). КОРРЕЛЯЦИОННЫЕ МЕТОДЫ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ ПУТЕМ ГОЛОСОВАНИЯ СИСТЕМ ФРАГМЕНТОВ. *Радіоелектроніка, інформатика, управління*, (1).

17. Павлова М. О. Анімація графічних об'єктів у технології захоплення рухів / М. О. Павлова // *Радіоелектроніка та молодь у ХХІ столітті: 23 міжнародний молодіжний форум*, 16-18 квітня 2019 р.–154 с.

18. Павлова М. О. Метод ланцюгового кодування Фрімена у комп'ютерному зорі / М. О. Павлова // *Радіоелектроніка та молодь у ХХІ столітті: 24 міжнародний молодіжний форум*, 7-9 квітня 2020 р.–18 с.

19. Дмитрий Лурье (2012). 3D, HD, 1080p, 1080i, Blue-Ray, DVD... Всё, что вы должны знать о видео // . URL: <http://neuezeiten.rusverlag.de/2012/01/18/1416-2>.

20. Коннова, Н. С., & Хаперская, В. Ю. (2020). Применение нейросетей и алгоритмов машинного обучения с целью диагностики сердечно-сосудистых заболеваний на основе сигналов СКГ. Биомедицинская радиоэлектроника, 23(1), 5-20.

21. F. Rosenblatt (1958). The perceptron: a probabilistic model for information storage and organization in the brain.

22. J. J. Hopfield (2006). Neural Networks and Physical Systems with Emergent Collective Computational Abilities.

23. G. E. Hinton, T. J. Sejnowski (1986). Learning and relearning in Boltzmann machines.

24. Paul Smolensky (1986). Information Processing in Dynamical Systems: Foundations of Harmony Theory.

25. H. Bourlard, Y. Kamp (2004). Auto-association by multilayer perceptrons and singular value decomposition.

26. The Scientific Research Honor Society (2020). The Hominin next door.

27. Lee, D.D. and Seung, H.S. (1999) Learning the parts of objects by non-negative matrix factorization. Nature, 401:788-791.

28. Hyvarinen, A. and Hoyer, P.O. (2001) A 2-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images. Vision Research, 41:2413-2423.

29. Olshausen, B.A. (2002) Sparse codes and spikes. R.P.N. Rao, B.A. Olshausen and M.S. Lewicki Eds. – MIT press:257-272.

30. Teh, Y.W. and Welling, M. and Osindero, S. and Hinton, G.E. (2003) Energy-based models for sparse overcomplete representations. Journal of Machine Learning Research, 4:1235-1260.

31. Стариков, А. (2010). Самоорганизующиеся карты Кохонена-математический аппарат. URL: <http://www.basegroup.ru/library/analysis/clusterization/som>.
32. Simoncelli, E.P. (2005) Statistical modeling of photographic images. Academic Press 2nd ed.
33. Hinton, G.E. and Zemel, R.S. (1994) Autoencoders, minimum description length, and Helmholtz free energy. Advances in Neural Information Processing Systems 6, J. D. Cowan, G. Tesauro and J. Alspector (Eds.), Morgan Kaufmann: San Mateo, CA.
34. Olshausen, B.A. and Field, D.J. (1997) Sparse coding with an overcomplete basis set: a strategy employed by V1? Vision Research, 37:3311-3325.
35. Foldiak, P. (1990) Forming sparse representations by local anti-hebbian learning. Biological Cybernetics, 64:165-170.
36. Park, Y. (2019). Image Segmentation by Cascaded Superpixel Merging with Privileged Information. Journal of the Korea Institute of Information and Communication Engineering, 23(9), 1049-1059. URL: <http://www.cs.berkeley.edu/projects/vision/grouping/segbench/>.
37. Simard, P.Y. Steinkraus, D. and Platt, J.C. (2003) Best practices for convolutional neural networks. ICDAR.
38. LeCun, Y. Bottou, L. Bengio, Y. and Haffner, P. (1998) Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11):2278-2324.
39. Hinton, G.E., Osindero, S. and Teh, Y. (2006) A fast learning algorithm for deep belief nets. Neural Computation 18, pp 1527-1554.
40. Diederik P. Kingma, Max Welling (2014). Auto-Encoding Variational Bayes.
41. Pascal Vincent, Hugo Larochelle, Yoshua Bengio, Pierre-Antoine Manzagol (2007). Extracting and Composing Robust Features with Denoising Autoencoders.

42. Oeldorf-Hirsch, A., & Sundar, S. S. (2015). Posting, commenting, and tagging: Effects of sharing news stories on Facebook. *Computers in human behavior*, 44, 240-249.

43. Daksh Trehan (2020). The inescapable AI algorithm: TikTok // Toward data science. URL: <https://towardsdatascience.com/the-inescapable-ai-algorithm-tiktok-ad4c6fd981b8>.

44. Janice Gassam Asare (2020). Does TikTok Have A Race Problem? // Forbs. URL: <https://www.forbes.com/sites/janicegassam/2020/04/14/does-tiktok-have-a-race-problem/?sh=65b3b09f3260>.

45. Traffic October 2020 Overview, TikTok // SimilarWeb. URL: <https://www.similarweb.com/website/tiktok.com>.

46. Tower, S. (2020). TikTok Crosses 2 Billion Downloads After Best Quarter For Any App Ever. URL: <https://sensortower.com/blog/tiktok-downloads-2-billion>.