

АНАЛІЗ ВІДТОКУ КОРИСТУВАЧІВ

У МОБІЛЬНОМУ ЗАСТОСУНКУ

МЕТА ДОСЛІДЖЕННЯ

Виявити поведінкові патерни користувачів
та визначити фактори, що впливають на відтік

- Порівняти поведінку різних сегментів
- Описати відмінності під час онбордингу та взаємодії
- Класифікувати користувачів із НА
- Запропонувати конкретні зміни в продукті

СТРУКТУРА ДАТАСЕТУ

#	user_id	install_date	install_time	event_counter	event_type	event_date	event_time
54	1	2025-04-15	2025-04-15 12:22:54+00:00	54	send_message_tap	2025-04-15	2025-04-15 15:48:01+00:00
55	1	2025-04-15	2025-04-15 12:22:54+00:00	55	answer_error	2025-04-15	2025-04-15 15:47:31+00:00
56	1	2025-04-15	2025-04-15 12:22:54+00:00	56	send_message_tap	2025-04-15	2025-04-15 15:48:01+00:00
57	1	2025-04-15	2025-04-15 12:22:54+00:00	57	answer_error	2025-04-15	2025-04-15 15:48:54+00:00
58	1	2025-04-15	2025-04-15 12:22:54+00:00	58	send_message_tap	2025-04-15	2025-04-15 15:49:40+00:00
59	1	2025-04-15	2025-04-15 12:22:54+00:00	59	answer_received	2025-04-15	2025-04-15 15:50:18+00:00
60	1	2025-04-15	2025-04-15 12:22:54+00:00	60	answer_dislike_tap	2025-04-15	2025-04-15 15:50:57+00:00
61	2	2025-04-25	2025-04-25 12:49:49+00:00	1	quiz_question_onboarding_screen_view	2025-04-25	2025-04-25 12:49:49+00:00
62	2	2025-04-25	2025-04-25 12:49:49+00:00	2	quiz_question_continue_tap	2025-04-25	2025-04-25 12:50:25+00:00
63	2	2025-04-25	2025-04-25 12:49:49+00:00	3	quiz_question_onboarding_screen_view	2025-04-25	2025-04-25 12:50:50+00:00
64	2	2025-04-25	2025-04-25 12:49:49+00:00	4	quiz_question_continue_tap	2025-04-25	2025-04-25 12:51:11+00:00
65	2	2025-04-25	2025-04-25 12:49:49+00:00	5	quiz_question_onboarding_screen_view	2025-04-25	2025-04-25 12:51:40+00:00
66	2	2025-04-25	2025-04-25 12:49:49+00:00	6	quiz_question_continue_tap	2025-04-25	2025-04-25 12:52:15+00:00
67	2	2025-04-25	2025-04-25 12:49:49+00:00	7	one_onboarding_screen_view	2025-04-25	2025-04-25 12:52:17+00:00
68	2	2025-04-25	2025-04-25 12:49:49+00:00	8	one_onboarding_continue_tap	2025-04-25	2025-04-25 12:52:19+00:00
69	2	2025-04-25	2025-04-25 12:49:49+00:00	9	two_onboarding_screen_view	2025-04-25	2025-04-25 12:52:20+00:00
70	2	2025-04-25	2025-04-25 12:49:49+00:00	10	two_onboarding_continue_tap	2025-04-25	2025-04-25 12:52:22+00:00
71	2	2025-04-25	2025-04-25 12:49:49+00:00	11	three_onboarding_screen_view	2025-04-25	2025-04-25 12:52:24+00:00
72	2	2025-04-25	2025-04-25 12:49:49+00:00	12	three_onboarding_continue_tap	2025-04-25	2025-04-25 12:52:26+00:00
73	2	2025-04-25	2025-04-25 12:49:49+00:00	13	subscription screen view	2025-04-25	2025-04-25 12:53:16+00:00

590,079 коректних записів

92,650 записів з **NA** статусом

13 невалідних записів

EXPLORATORY DATA ANALYSIS

Питання для дослідження:

1. Які характеристики описують поведінку користувачів?
2. Чи є відмінності між `retained` та `churned`?
3. Чи є зв'язок між діями та відтоком?
4. Як попередня поведінка впливає на рішення?

1.1 ДЕСКРИПТИВНІ ХАРАКТЕРИСТИКИ

Активність використання

- messages_sent, messages_received
- chat_views, total_events

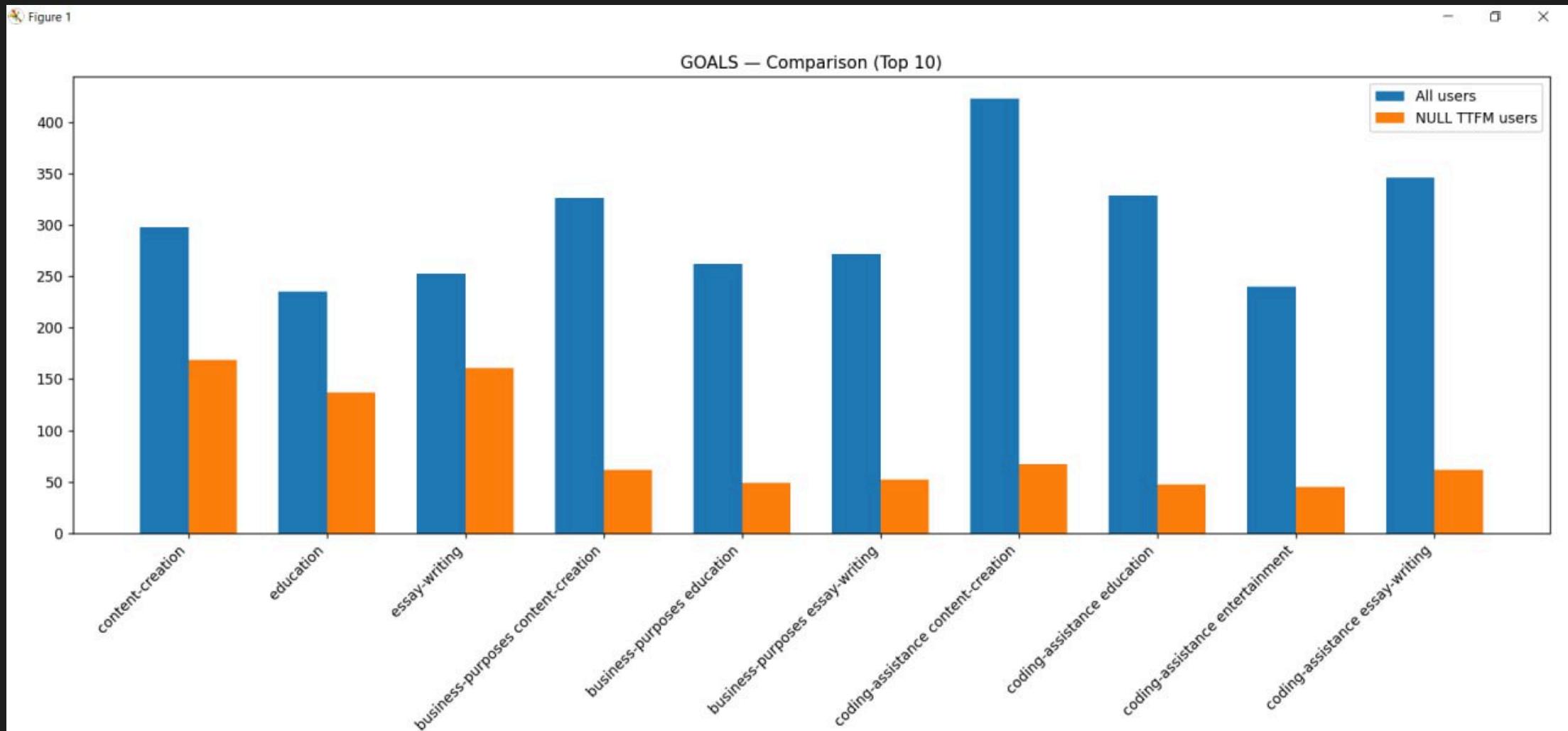
Оцінка задоволеності

- likes, dislikes, Like/Dislike ratio

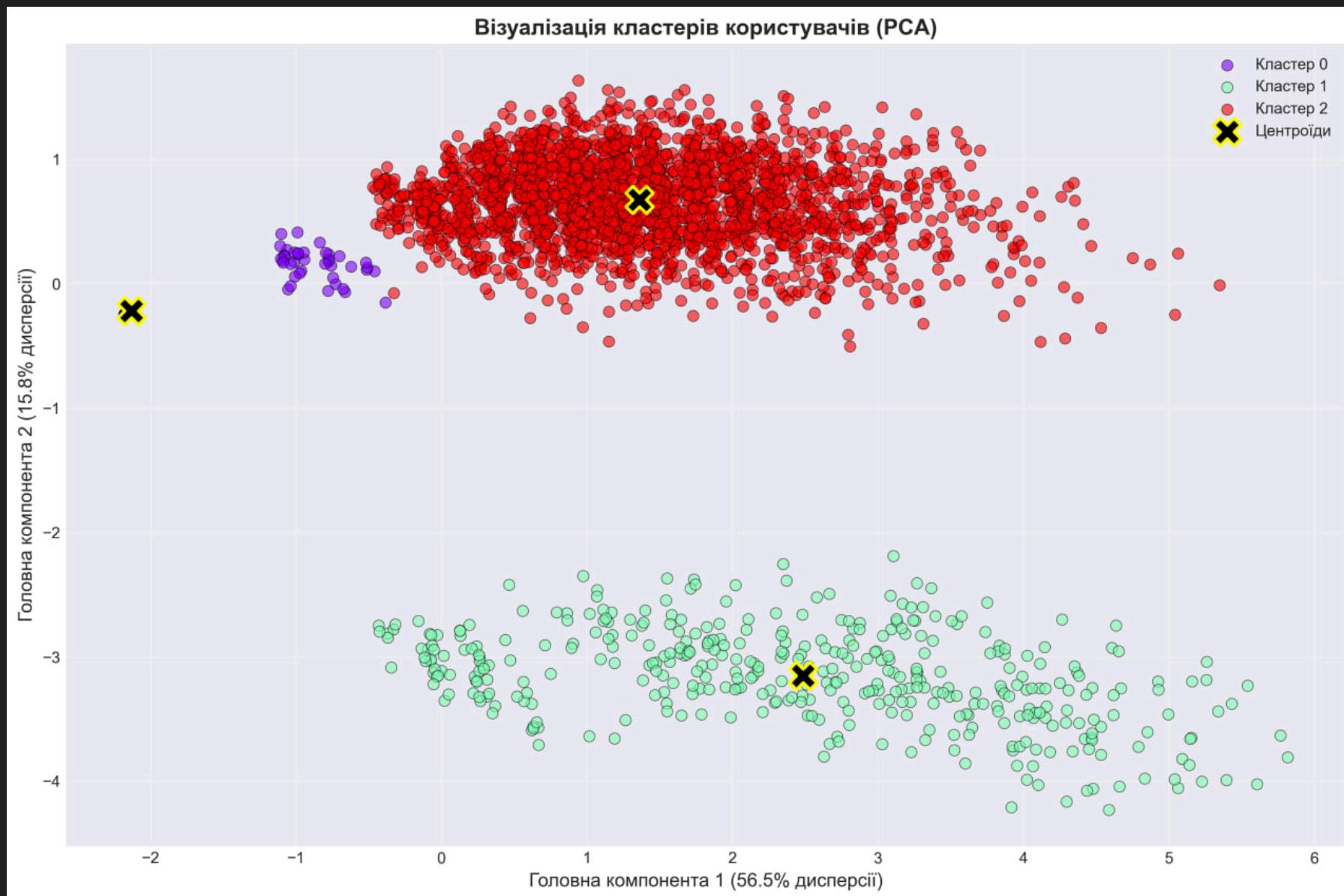
Якість сервісу

- error_rate

1.1 Відтік на етапі онбордингу



1.2 КЛАСТЕРИЗАЦІЯ КОРИСТУВАЧІВ



1.2 ХАРАКТЕРИСТИКИ КЛАСТЕРІВ

Кластер 0

Значно нижча активність за інші

Кластер 1

Значно менший відтік

Кластер 2

Значно більше помилок

1.2 ВІДМІННОСТІ МІЖ СЕГМЕНТАМИ

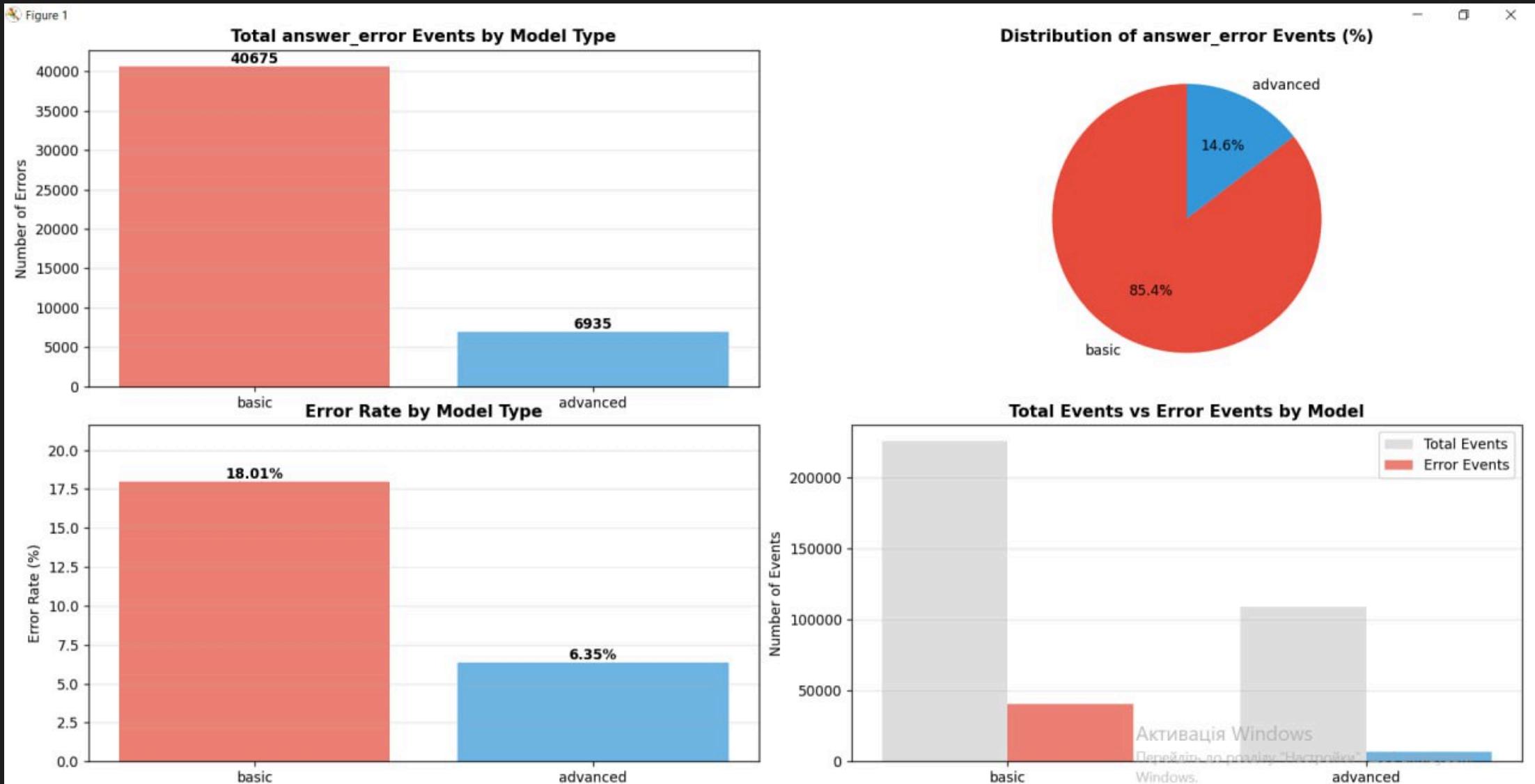
Retained користувачі:

- Висока активність
- Відносно низький error_rate

Churned користувачі:

- Дуже низький error_rate (ранній відтік)
- Або дуже високий error_rate (пізній відтік)

1.3 ВПЛИВ МОДЕЛІ НА ВІДТІК



1.3 ВПЛИВ МОДЕЛІ НА ВІДТІК

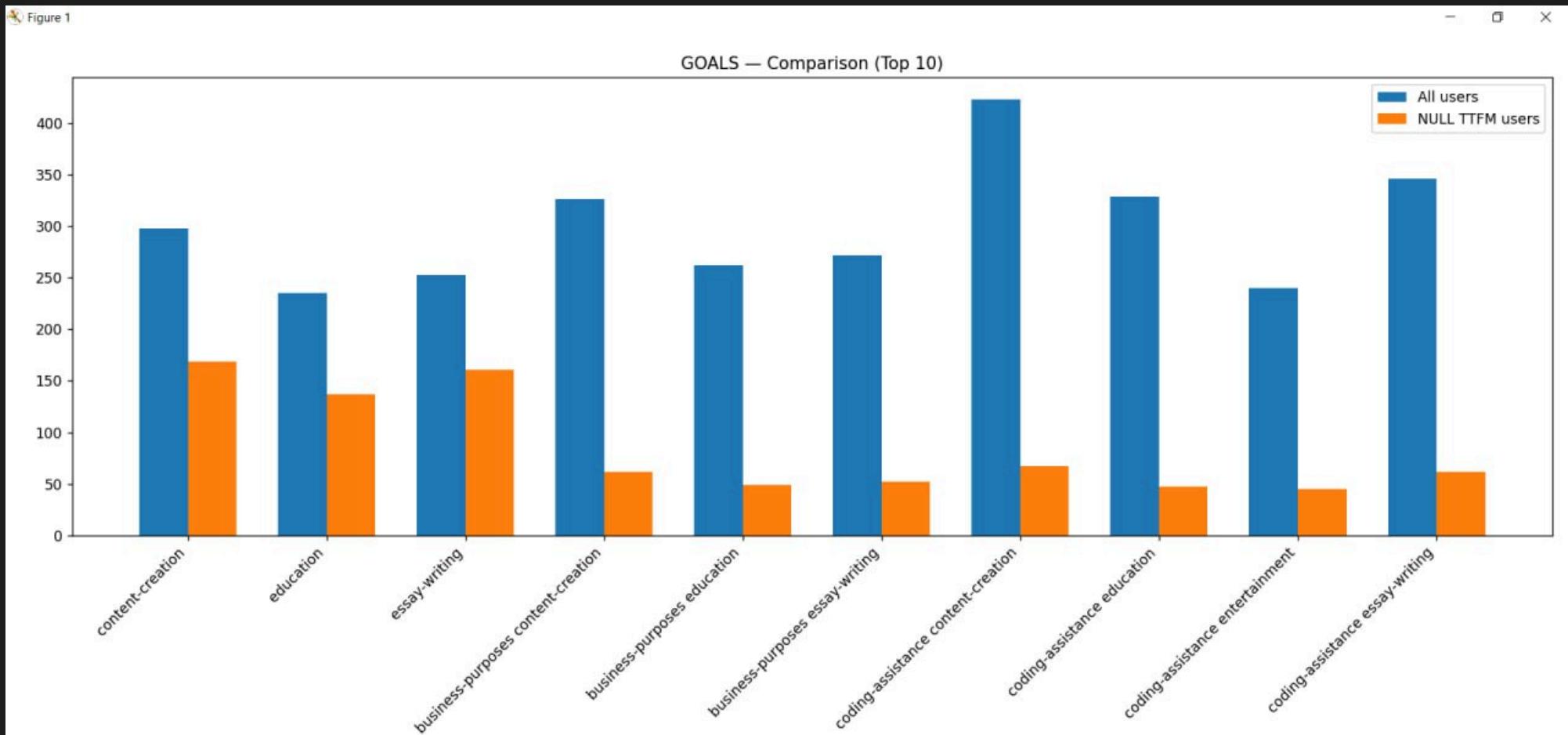
Базова модель:

- Retention rate: ~13%
- Високий error_rate

Advanced модель:

- Менше користувачів
- Значно менше помилок
- Вищий retention

1.3 ПРОБЛЕМНІ ІНТЕРЕСИ



Користувачі з інтересами **content-creation**, **education**, **essay-writing** швидше відпадають навіть не купивши

1.4 ФАКТОРИ ВПЛИВУ НА ВІДТІК

Критичні фактори:

1. Використання моделі

Для базової набагато більший відтік

2. Отримання помилок

Високий `error_rate` призводить до `churn`

3. Довге використання

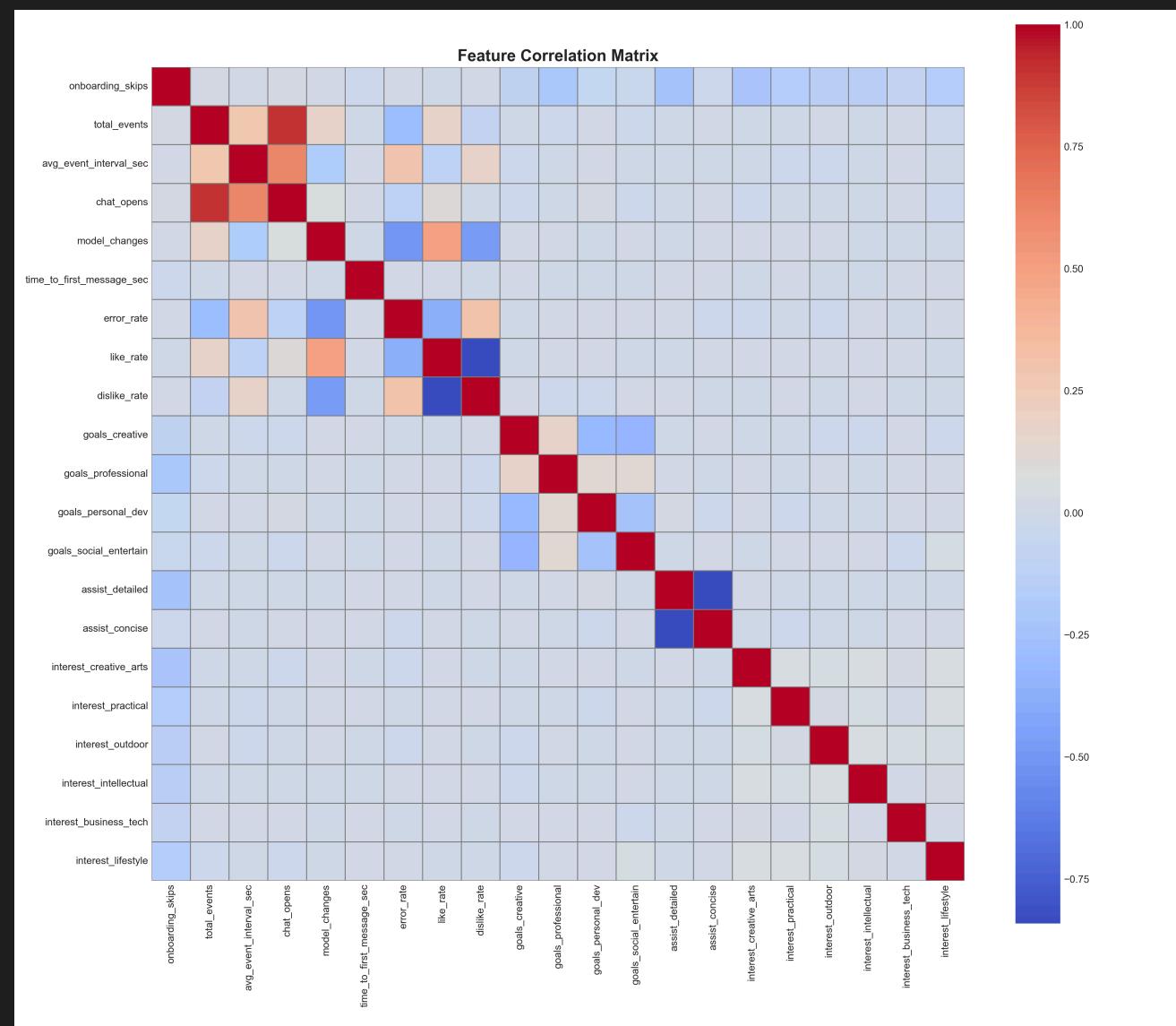
Чим довше час користування застосунком, тим менший шанс відтоку

1.4 ОБГРУНТУВАННЯ ВИБОРУ ФАКТОРІВ

Для кластеризації вибрали компоненти з найбільшою кореляцією з `is_churned`:

- `answer_errors`
- `avg_event_interval_sec`
- `chat_opens`, `chat_views`
- `dislikes`, `likes`
- `error_rate`
- `messages_received`, `messages_sent`
- `total_events`

1.5 КОРЕЛЯЦІЙНА МАТРИЦЯ



ЗАГАЛЬНА ІНФОРМАЦІЯ

3 класи користувачів:

1. Ранній відтік

Покидають застосунок зразу після онбордингу, відмовившись від преміум - підписки

2. Пізній відтік

Відтікають після тривалого використання, внаслідок високого error_rate

3. Активні користувачі

Не відтікають, мають низький error_rate

ПРИЧИНИ ВІДТОКУ

Виявлені проблеми:

1. Висока кількість помилок базової версії
error_rate ~9% для churned vs ~3% для retained
2. Невідповідність інтересам
content-creation, education, essay-writing
3. Бар'єр монетизації
Відмова від преміум після онбордингу

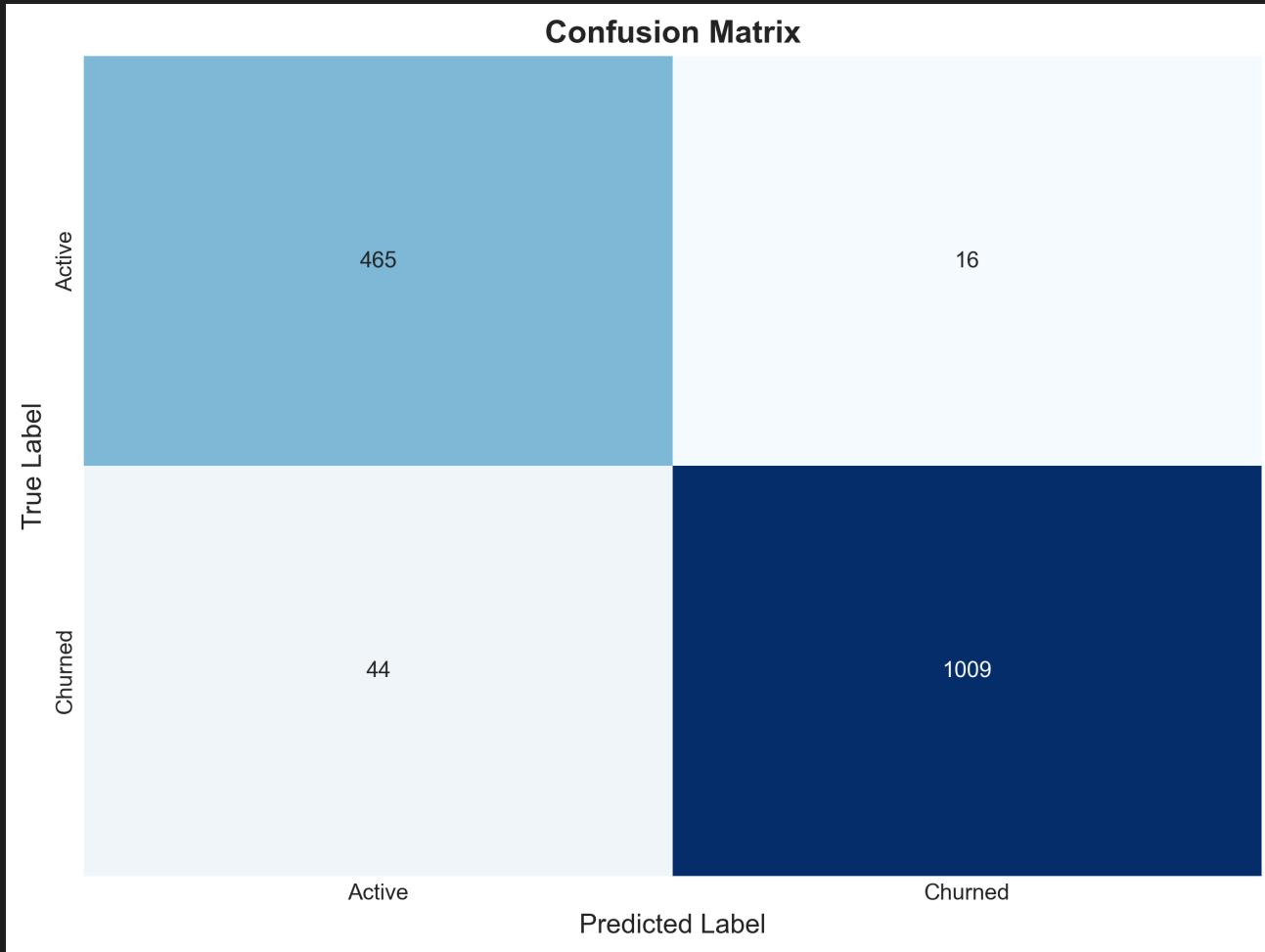
2. PREDICTIVE MODELING

Розроблені моделі:

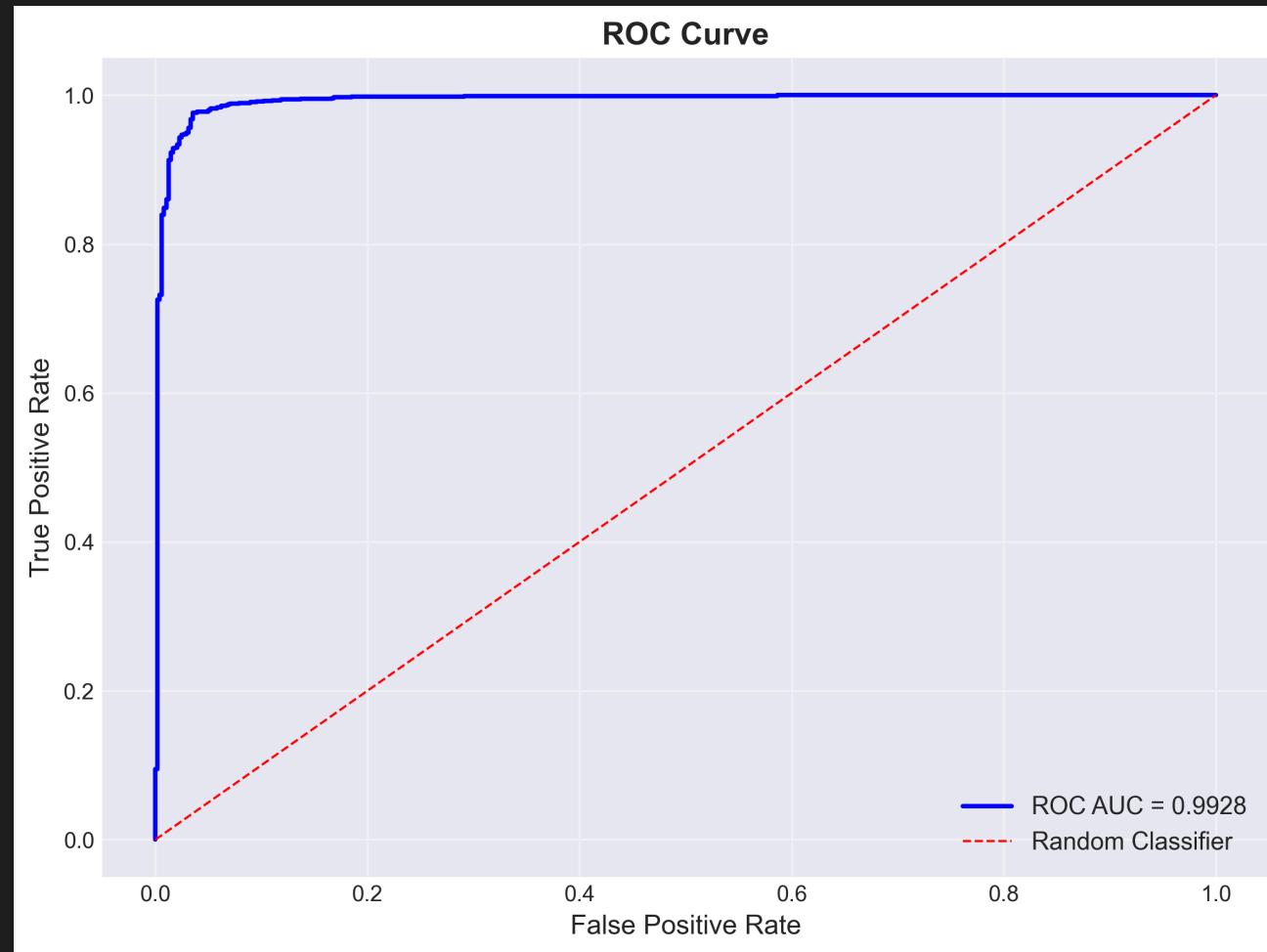
- Логістична регресія
- Random Forest

Мета: спрогнозувати чи користувач скасує підписку

2. ХАРАКТЕРИСТИКИ МОДЕЛІ



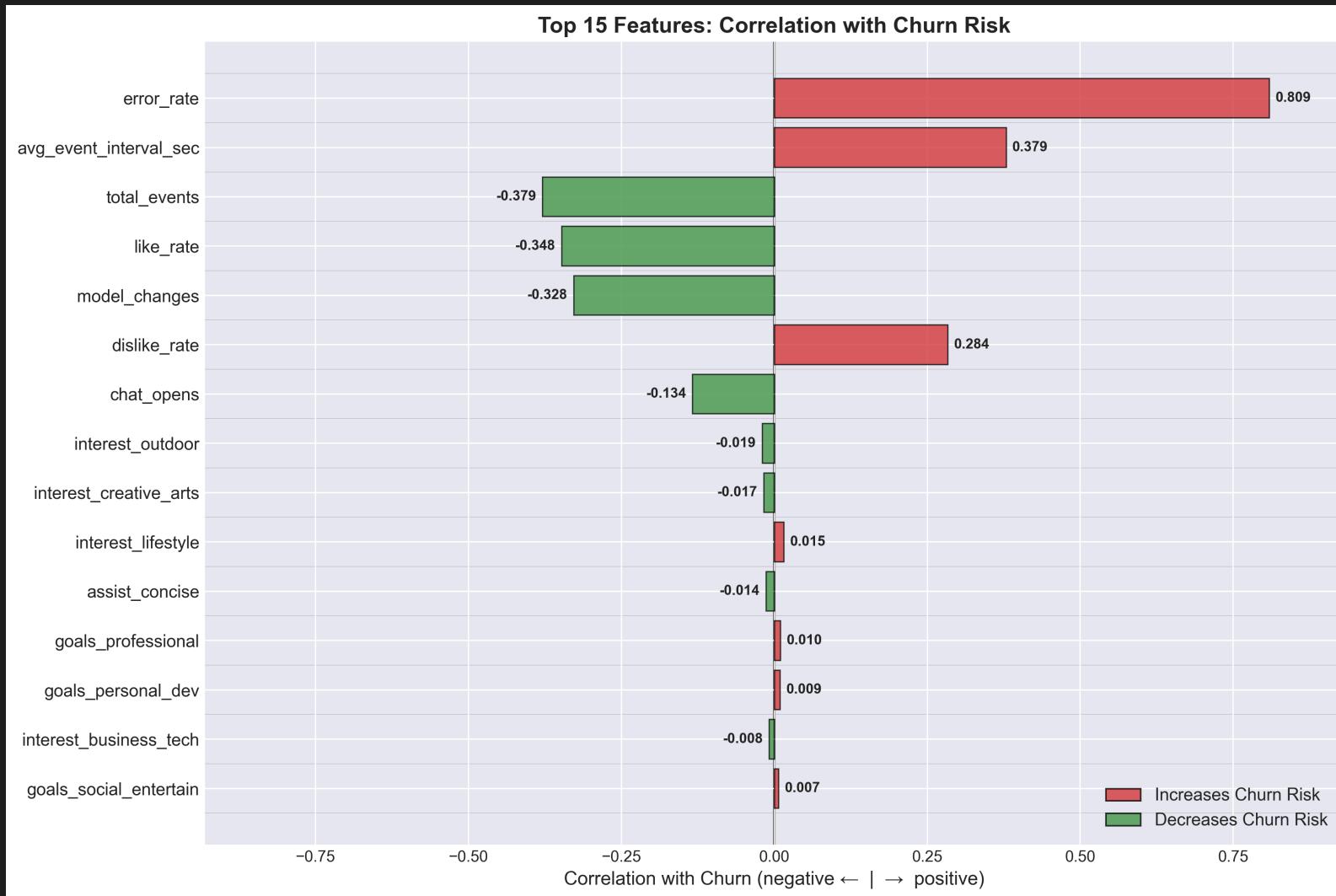
2. МЕТРИКИ ЕФЕКТИВНОСТИ



2. МЕТРИКИ ЕФЕКТИВНОСТИ

ACCURACY.....	0.9640	(± 0.0043)
PRECISION.....	0.9835	(± 0.0032)
RECALL.....	0.9636	(± 0.0043)
F1.....	0.9735	(± 0.0031)
ROC_AUC.....	0.9934	(± 0.0018)

2. FEATURE IMPORTANCE



Найважливіші фактори для прогнозування відтоку

РЕКОМЕНДАЦІЇ: ТЕХНІЧНІ

1. Покращити стабільність

- Базова модель: **критично**
- Преміум модель: бажано
- Таргет: `error_rate < 5%`

2. Промо **advanced** моделі

- Наголосити на перевагах відносно базової версії
- Таргет: базові користувачі

РЕКОМЕНДАЦІЇ: ПРОДУКТОВІ

Retention для проблемних інтересів:

- Розробити спеціальні фічі для:
 - content-creation
 - education
 - essay-writing
- Запропонувати при онбордингу
- Провести АВ-тестування
- Додати нові опції

ПІДСУМОК

Проаналізовано: 590,079 подій, 10,218 користувачів

Виявлено: 3 кластери поведінки

Критичні проблеми:

- Високий error_rate базової моделі
- Низький retention специфічних інтересів

Retention rate базової моделі: ~13%

NEXT STEPS

Пріоритети:

1. Покращити стабільність базової моделі
2. Розробити спеціальні фічі
3. АВ-тестування нових підходів
4. Промо advanced моделі
5. Оптимізація paywall

ДЯКУЮ ЗА УВАГУ!