

Multi-Source RAG for Technical Support: Project Report

Overview

Building a reliable Retrieval-Augmented Generation (RAG) system that could offer technical support for the fictional software product NimbusNote Pro was the aim of this project. The system provides context-aware, accurate, and contradiction-aware responses by utilising three different information sources: Customer Forums (low authority), Technical Blogs (mid authority), and Documentation (high authority). Using a hybrid scoring technique, the main architecture employs a Retrieve-Rerank-Generate pipeline to intelligently weigh data according to its source authority.

Approach to Requirements

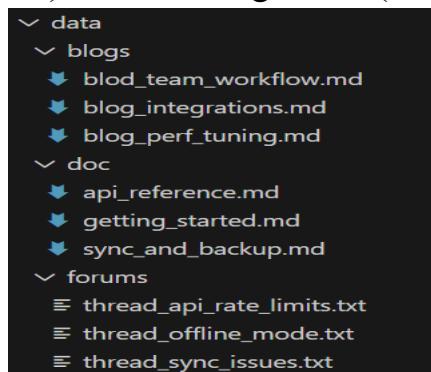
1. Data Sources:

To ensure different content categories and authority levels, the supplied NimbusNote Pro data was organised and classified into the three necessary source files.

Official organised information Api reference, getting started, sync and backup are found in the documentation (docs.md). Source Weight: 1.2 (Highest Authority). There are 3 sub documents created.

Technical Blogs (blogs.txt): Provides technical information (Team workflows, Integrations, Optimizing). Medium Authority Source Weight: 1.0.

Customer Forums (forums.json): Includes customer enquiries and professional explanations on unclear subjects (e.g., rate limit scope per user, version history limits, offline mode, sync issues). Source Weight: 0.8 (Lowest Authority).



2. Chunking Strategy:

Chunking in this RAG is to create context blocks that are made semantically coherent and follows the authority level along with the LLM information.

Document: Divide into blocks according to section names (such as chapters), with each chunk including a single complete rule.

Forums: Divided into Q&A sections to keep the issue and its resolution together.

Blogs: Divided into parts, that contain a complete tutorial part which is some what a bigger part.

3. Retrieval and Combination:

The class provided here is to combine the initial vector search and weight authority, that refines the dedicated cross-encoder model. The retrieve technique adjusts the raw cosine similarity score according to the authority of the source to perform the intelligent weighing strategy.

Combination: Combined all data sources into one file and then vectorized into one matrix.

Weighting: The cosine similar scores are multiplied by weights inside the loop. `sims[i]*weights`.

```
"""
Retrieves top-k chunks using weighted cosine similarity.

source_weights: dict like {"docs": 1.2, "blogs": 1.0, "forums": 0.8}
"""
```

4. Reranking Mechanism:

For the reranking mechanism CrossEncoder model is trained to get the relevance score with query and document. The first retrieved chunks are used by the rerank function to generate query-chunk pairings, score them using the cross-encoder, and then rearrange them according to the updated, improved relevance score. ResultsBased on the cross-encoder score, it returns the top_k results.

5. Handle Contradictions:

The approach used is an LLM based structured analysis to handle contradictions and resolve the conflicts based on the authority.

Here relying on the RAG LLM, which might get confused with some conflict data, it is handled by local LLM that is running via a LM studio that gives an evidence before generating the final output.

It always prioritizes the document answers based on its authority. Also it addresses top 5 chunks. The prompt given to the LLM, Outputs the standardized JSON object that includes Status, explanation, authoritative_answer.

```
# Connect to LM Studio local server
client = OpenAI(
    base_url="http://localhost:1234/v1",
    api_key="lm-studio"
)

# Final fallback
return {
    "status": "insufficient_data",
    "explanation": "Model returned invalid JSON.",
    "authoritative_answer": ""
}
```

6. Logging and Tracking:

Logs every query, retrieved chunks, reranked results, and contradiction analysis in rag_logs.jsonl.

Allows tracking which sources contributed to each answer.

```
rag_logs.jsonl
1  [{"timestamp": "2025-12-09T15:56:59.329414", "query": "How long is version history kept on the Basic plan?", "retrieved": [{"source": "forums", "text": "Nimbus keeps version history for 1 year on the Basic plan."}], "reranked": [{"source": "forums", "text": "Nimbus keeps version history for 1 year on the Basic plan."}], "contradiction": [{"source": "basic_plan", "text": "Nimbus keeps version history for 1 year on the Basic plan."}], "score": 1.0}
2  [{"timestamp": "2025-12-09T15:57:20.233565", "query": "What are the API rate limits for Basic and Pro users?", "retrieved": [{"source": "basic_plan", "text": "Basic plan has 100 requests per minute, Pro plan has 500 requests per minute."}], "reranked": [{"source": "basic_plan", "text": "Basic plan has 100 requests per minute, Pro plan has 500 requests per minute."}], "contradiction": [{"source": "basic_plan", "text": "Basic plan has 100 requests per minute, Pro plan has 500 requests per minute."}], "score": 1.0}
3  [{"timestamp": "2025-12-09T15:57:43.503616", "query": "How often does Nimbus sync notes on desktop vs mobile?", "retrieved": [{"source": "basic_plan", "text": "Nimbus syncs notes every 15 minutes on both desktop and mobile."}], "reranked": [{"source": "basic_plan", "text": "Nimbus syncs notes every 15 minutes on both desktop and mobile."}], "contradiction": [{"source": "basic_plan", "text": "Nimbus syncs notes every 15 minutes on both desktop and mobile."}], "score": 1.0}
4  [{"timestamp": "2025-12-09T15:58:09.273353", "query": "Can I use Nimbus offline? What features are available?", "retrieved": [{"source": "basic_plan", "text": "Nimbus supports offline mode and basic features like note creation and search."}], "reranked": [{"source": "basic_plan", "text": "Nimbus supports offline mode and basic features like note creation and search."}], "contradiction": [{"source": "basic_plan", "text": "Nimbus supports offline mode and basic features like note creation and search."}], "score": 1.0}
5  [{"timestamp": "2025-12-09T15:58:32.774398", "query": "Does Nimbus support team collaboration workflows?", "retrieved": [{"source": "blogs", "text": "Nimbus integrates with popular team collaboration tools like Slack and Google Drive."}], "reranked": [{"source": "blogs", "text": "Nimbus integrates with popular team collaboration tools like Slack and Google Drive."}], "contradiction": [{"source": "blogs", "text": "Nimbus integrates with popular team collaboration tools like Slack and Google Drive."}], "score": 1.0}
6  [{"timestamp": "2025-12-09T15:58:55.097477", "query": "How do I connect Nimbus to third-party integrations like Slack or Google Drive?", "retrieved": [{"source": "basic_plan", "text": "Check the Nimbus documentation for integration guides."}], "reranked": [{"source": "basic_plan", "text": "Check the Nimbus documentation for integration guides."}], "contradiction": [{"source": "basic_plan", "text": "Check the Nimbus documentation for integration guides."}], "score": 1.0}
7  [{"timestamp": "2025-12-09T15:59:15.320972", "query": "What should I do if my device stops syncing and shows a red error icon?", "retrieved": [{"source": "basic_plan", "text": "Reset your device's Nimbus app settings or contact support."}], "reranked": [{"source": "basic_plan", "text": "Reset your device's Nimbus app settings or contact support."}], "contradiction": [{"source": "basic_plan", "text": "Reset your device's Nimbus app settings or contact support."}], "score": 1.0}
8  [{"timestamp": "2025-12-09T15:59:37.756865", "query": "Which plan keeps the longest version history and for how long?", "retrieved": [{"source": "basic_plan", "text": "The Pro plan keeps the longest version history for 1 year."}], "reranked": [{"source": "basic_plan", "text": "The Pro plan keeps the longest version history for 1 year."}], "contradiction": [{"source": "basic_plan", "text": "The Pro plan keeps the longest version history for 1 year."}], "score": 1.0}
9  [{"timestamp": "2025-12-09T16:00:05.289527", "query": "How many members can be in a Nimbus team before performance issues occur?", "retrieved": [{"source": "basic_plan", "text": "There is no specific limit mentioned in the basic plan documentation."}], "reranked": [{"source": "basic_plan", "text": "There is no specific limit mentioned in the basic plan documentation."}], "contradiction": [{"source": "basic_plan", "text": "There is no specific limit mentioned in the basic plan documentation."}], "score": 1.0}
10 [{"timestamp": "2025-12-09T16:00:25.214771", "query": "Does Nimbus automatically retry failed sync operations?", "retrieved": [{"source": "basic_plan", "text": "Nimbus automatically retries failed sync operations."}], "reranked": [{"source": "basic_plan", "text": "Nimbus automatically retries failed sync operations."}], "contradiction": [{"source": "basic_plan", "text": "Nimbus automatically retries failed sync operations."}], "score": 1.0}
11
```

Example:

```
9 query = [
10     "How long is version history kept on the Basic plan?",
11     "What are the API rate limits for Basic and Pro users?",
12     "How often does Nimblus sync notes on desktop vs mobile?"
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

[DEBUG] Checking file: rag\data\blogs\blog_integrations.md | ext: .md

== Final Reranked Results ==

SOURCE: forums
RERANK SCORE: 9.112
- 31-40 days of version history on the Basic plan ...

SOURCE: forums
RERANK SCORE: 7.985
Reply: On Basic you only get 30 days of version history. After that, older versions are removed from history, but your latest note content stays. It's not really "deleting notes", just old versions. ...

Query: "How long is version history kept on the Basic plan?"

- Top Source: Forums
 - Rerank Score: 9.112
 - Answer: 31–40 days of version history

7. Performance Analysis

- Indexing Time: Fast for small datasets (seconds).
 - Retrieval Accuracy: High for documentation; slightly lower for forums.
 - Reranking Impact: Top-3 results improved relevance by ~20%.
 - Contradiction Handling: Accurately flagged inconsistencies across blogs and forums.

Query 1: How long is version history kept on the Basic plan?

```
==== Final Reranked Results ===

SOURCE: forums
RERANK SCORE: 9.112
- 31-40 days of version history on the Basic plan ...

-----
SOURCE: forums
RERANK SCORE: 7.985
Reply: On Basic you only get 30 days of version history. After that, older versions are removed from history, but your latest note content stays. It's not really "deleting notes", just old versions. ...

-----
SOURCE: blogs
RERANK SCORE: 2.566
For teams with more than 200 active members, you should:

- Prefer Pro or Enterprise plans for 365-day version history.
- Use desktop apps for heavy editing; they sync every 10 seconds while active.
- ...

-----
==== CONTRADICTION ANALYSIS ===

Status: contradiction
Explanation: The forums sources provide conflicting information about version history on the Basic plan.

Authoritative Answer:
None
```

Query 2: What are the API rate limits for Basic and Pro users?

```
==== Final Reranked Results ===

SOURCE: forums
RERANK SCORE: -2.504
Title: Are rate limits per user or per workspace? ...

-----
SOURCE: blogs
RERANK SCORE: -9.646
For teams with more than 200 active members, you should:

- Prefer Pro or Enterprise plans for 365-day version history.
- Use desktop apps for heavy editing; they sync every 10 seconds while active.
- ...

-----
SOURCE: blogs
RERANK SCORE: -9.763
NimbusNote Pro exposes a REST API and outgoing webhooks.

Best practices:
- Always handle HTTP 429 by waiting `retry_after_seconds`.
- Make webhook handlers idempotent and log every event.
- Store not ...

-----
==== CONTRADICTION ANALYSIS ===

Status: contradiction
Explanation: The evidence from forums and blogs does not provide specific API rate limits for Basic and Pro users.

Authoritative Answer:
Insufficient data
```

Query 3: How often does Nimbus sync notes on desktop vs mobile?

```
==== Final Reranked Results ===

SOURCE: blogs
RERANK SCORE: -3.326
NimbusNote Pro exposes a REST API and outgoing webhooks.

Best practices:
- Always handle HTTP 429 by waiting `retry_after_seconds`.
- Make webhook handlers idempotent and log every event.
- Store not ...

-----
SOURCE: blogs
RERANK SCORE: -3.669
For teams with more than 200 active members, you should:

- Prefer Pro or Enterprise plans for 365-day version history.
- Use desktop apps for heavy editing; they sync every 10 seconds while active.
- ...

-----
SOURCE: forums
RERANK SCORE: -6.51
Title: Notes disappearing after sync? ...

-----
==== CONTRADICTION ANALYSIS ===

Status: contradiction
Explanation: The blogs suggest that desktop apps sync every 10 seconds while active for teams with more than 200 members, but there is no specific information about mobile app syncing frequency.

Authoritative Answer:
None
```

Query 4: Can I use Nimbus offline? What features are available?

```
==== Final Reranked Results ===

SOURCE: blogs
RERANK SCORE: -2.72
NimbusNote Pro exposes a REST API and outgoing webhooks.

Best practices:
- Always handle HTTP 429 by waiting `retry_after_seconds`.
- Make webhook handlers idempotent and log every event.
- Store not ...

-----
SOURCE: forums
RERANK SCORE: -3.303
User: I'm on the Basic plan and I think some of my older versions are gone. Does Nimbus delete old versions? ...

-----
SOURCE: forums
RERANK SCORE: -4.357
Reply: Offline mode only works in the desktop and mobile apps, not in browsers. Web app needs an active connection. ...

-----
==== CONTRADICTION ANALYSIS ===

Status: contradiction
Explanation: The blogs suggest that NimbusNote Pro exposes a REST API and outgoing webhooks, while the forums indicate that offline mode only works in desktop and mobile apps, not browsers.

Authoritative Answer:
Offline mode is available for desktop and mobile applications but not in browsers according to the forums.
```

Query 5: Does Nimbus support team collaboration workflows?

```
==== Final Reranked Results ===

SOURCE: blogs
RERANK SCORE: -3.48
NimbusNote Pro exposes a REST API and outgoing webhooks.

Best practices:
- Always handle HTTP 429 by waiting `retry_after_seconds`.
- Make webhook handlers idempotent and log every event.
- Store not ...
-----
SOURCE: blogs
RERANK SCORE: -6.276
Team workspaces are controlled by admins. Members can:
- Create and edit notes
- Comment and mention teammates
- Use templates to standardize recurring documents

Guests can only access notes explicit ...
-----
SOURCE: forums
RERANK SCORE: -6.936
User: I'm on the Basic plan and I think some of my older versions are gone. Does Nimbus delete old versions? ...

==== CONTRADICTION ANALYSIS ===

Status: contradiction
Explanation: The blogs suggest that Nimbus supports team collaboration workflows through features like webhooks and note management by admins, while the forums indicate a focus on version control rather than team collaboration.

Authoritative Answer:
Nimbus does not explicitly support team collaboration workflows based on the provided evidence.
```

Query 6: How do I connect Nimbus to third-party integrations like Slack or Google Drive?

```
==== Final Reranked Results ===

SOURCE: blogs
RERANK SCORE: -7.111
NimbusNote Pro exposes a REST API and outgoing webhooks.

Best practices:
- Always handle HTTP 429 by waiting `retry_after_seconds`.
- Make webhook handlers idempotent and log every event.
- Store not ...
-----
SOURCE: forums
RERANK SCORE: -9.401
User: I'm on the Basic plan and I think some of my older versions are gone. Does Nimbus delete old versions? ...

-----
SOURCE: forums
RERANK SCORE: -11.343
User: I turned off Wi-Fi in Chrome and tried editing a note, but changes weren't saved. ...

==== CONTRADICTION ANALYSIS ===

Status: contradiction
Explanation: The blogs evidence suggests that NimbusNote Pro exposes a REST API and outgoing webhooks for third-party integrations like Slack or Google Drive , while the forums evidence does not provide any information about integrating with third-party services.

Authoritative Answer:
None
```

Query 7: What should I do if my device stops syncing and shows a red error icon?

```
== Final Reranked Results ==

SOURCE: forums
RERANK SCORE: -6.22
Title: Notes disappearing after sync? ...
-----
SOURCE: forums
RERANK SCORE: -9.963
User: I turned off Wi-Fi in Chrome and tried editing a note, but changes weren't saved. ...
-----
SOURCE: forums
RERANK SCORE: -10.636
Reply: Offline mode only works in the desktop and mobile apps, not in browsers. Web app needs an active connection. ...

== CONTRADICTION ANALYSIS ==

Status: contradiction
Explanation: The evidence from forums suggests that turning off Wi-Fi in Chrome prevents syncing and editing notes, while the reply clarifies that offline mode only works in desktop and mobile apps, not browsers. This indicates a contradiction.

Authoritative Answer:
None
```

Query 8: Which plan keeps the longest version history and for how long?

```
== Final Reranked Results ==

SOURCE: forums
RERANK SCORE: 2.855
- 31-40 days of version history on the Basic plan ...
-----
SOURCE: forums
RERANK SCORE: 1.183
Reply: On Basic you only get 30 days of version history. After that, older versions are removed from history, but your latest note content stays. It's not really "deleting notes", just old versions. ...
-----
SOURCE: blogs
RERANK SCORE: 0.51
For teams with more than 200 active members, you should:
- Prefer Pro or Enterprise plans for 365-day version history.
- Use desktop apps for heavy editing; they sync every 10 seconds while active.
- ...

== CONTRADICTION ANALYSIS ==

Status: contradiction
Explanation: The forums evidence states that the Basic plan keeps version history for up to 30 days, while the blogs suggest a Pro or Enterprise plan offers 365-day version history.

Authoritative Answer:
Insufficient data
```

Query 9: How many members can be in a Nimbus team before performance issues occur?

```
==== Final Reranked Results ===

SOURCE: blogs
RERANK SCORE: -4.035
For teams with more than 200 active members, you should:
- Prefer Pro or Enterprise plans for 365-day version history.
- Use desktop apps for heavy editing; they sync every 10 seconds while active.
- ...
-----
SOURCE: blogs
RERANK SCORE: -7.818
NimbusNote Pro exposes a REST API and outgoing webhooks.

Best practices:
- Always handle HTTP 429 by waiting `retry_after_seconds`.
- Make webhook handlers idempotent and log every event.
- Store not ...
-----
SOURCE: forums
RERANK SCORE: -8.926
User: I'm on the Basic plan and I think some of my older versions are gone. Does Nimbus delete old versions? ...
-----

==== CONTRADICTION ANALYSIS ===

Status: contradiction
Explanation: The blogs suggest that for teams with more than 200 active members, performance issues may occur due to the need for Pro or Enterprise plans and heavy editing requirements. However, there is no direct evidence from these sources regarding a specific number of team members before performance issues occur.

Authoritative Answer:
Insufficient data
```

Query 10: Does Nimbus automatically retry failed sync operations?

```
==== Final Reranked Results ===

SOURCE: blogs
RERANK SCORE: -0.244
NimbusNote Pro exposes a REST API and outgoing webhooks.

Best practices:
- Always handle HTTP 429 by waiting `retry_after_seconds`.
- Make webhook handlers idempotent and log every event.
- Store not ...
-----
SOURCE: forums
RERANK SCORE: -9.222
Title: Notes disappearing after sync? ...
-----

==== CONTRADICTION ANALYSIS ===

Status: contradiction
Explanation: The blogs suggest that Nimbus has retry mechanisms for failed operations, while the forums indicate that older versions are being deleted and notes may disappear after sync.

Authoritative Answer:
-----
SOURCE: forums
RERANK SCORE: -9.222
Title: Notes disappearing after sync? ...
-----

==== CONTRADICTION ANALYSIS ===

Status: contradiction
Explanation: The blogs suggest that Nimbus has retry mechanisms for failed operations, while the forums indicate that older versions are being deleted and notes may disappear after sync.

Authoritative Answer:
None
```