

Retrieval-Augmented Generation (RAG) Model Development for Insurance Document Processing (HelpMate AI)

Background

The insurance industry deals with a vast amount of textual data, including policies, claims, legal documents, and customer inquiries. Traditional search-based retrieval methods often fail to provide precise and contextually relevant answers. Large Language Models (LLMs) can generate responses, but they struggle with factual accuracy when relying solely on their training data. Retrieval-Augmented Generation (RAG) bridges this gap by integrating an external knowledge source for accurate and contextually aware responses.

Problem Statement

Insurance professionals and customers frequently need precise answers from extensive insurance documentation. The challenge lies in:

- Extracting accurate and relevant information from large document repositories.
- Enhancing the response accuracy of LLMs with real-time document retrieval.
- Ensuring the retrieval process is efficient, reducing latency in responses.

A RAG-based system offers an effective solution by combining vector-based retrieval with generative AI, ensuring precise, contextual, and up-to-date responses.

Approach

The proposed solution leverages a RAG-based architecture to enhance insurance document processing. The key steps involved in the system are:

1. Building the Vector Store:

- Insurance documents are converted into vector representations using an embedding model.
- These vectors are stored in a vector database for efficient retrieval.

2. Cache, Search, and Re-rank:

- Incoming queries are checked against a cache to determine if similar queries have been processed previously.
- If not found in the cache, the query undergoes a vector search to retrieve the most relevant document chunks.

- Cross-encoders re-rank the top-k retrieved documents to enhance result precision.

3. **Generative Search:**

- The top-ranked documents are fed into an LLM along with the query.
- The LLM generates a contextually aware and precise response.

System Layers

The RAG system is structured into the following layers:

1. **Data Processing Layer:**

- Document ingestion and preprocessing.
- Text chunking and embedding generation using an embedding model.
- Storage in a vector database.

2. **Retrieval Layer:**

- Query processing and cache lookup.
- Vector search in the database.
- Top-k document selection based on similarity scores.

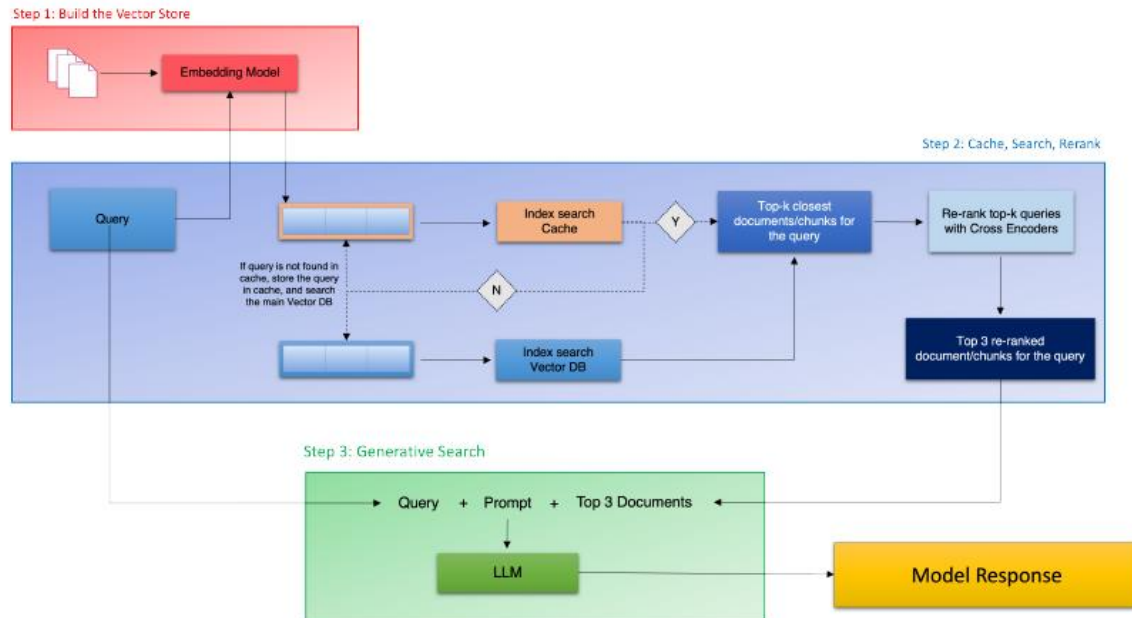
3. **Re-ranking Layer:**

- Cross-encoders refine the top-k results.
- Selection of the most relevant documents for LLM input.

4. **Generation Layer:**

- LLM processes the retrieved documents.
- Generates an informed response based on retrieved knowledge.
- Presents the final response to the user.

Architecture Diagram



Conclusion

The RAG-based model enhances the accuracy and efficiency of information retrieval in insurance document processing. By integrating vector search with generative AI, the system ensures precise, contextually relevant, and up-to-date responses. This approach significantly improves user experience and decision-making in the insurance domain, making complex information easily accessible.

Screenshots

1. Query – “is dental treatment included?”

Index	Metadatas	Documents	Distances	IDs	Reranked_scores
	[Page No.: 'Page 17', Policy Name: 'HDFC Life-Easy-Health-101N110V03-Policy-Bond-Single-Pay']	7. Routine eye tests, any Dental Treatment or Surgery of cosmetic nature, extraction of impacted tooth/teeth, orthodontics or orthognathic surgery, or tempero-mandibular joint disorder except as necessitated by an accidental injury and warranting Hospitalization. 8. Outpatient treatment. 9. Hospitalization and/or Surgery relating to infertility or impotency, sex change or any treatment related to it, abortion, sterilization and contraception including any complications relating thereto. 10. Hospitalization and/or Surgery for treatment arising from pregnancy and its complications which shall include childbirth or miscarriage. 11. Hospitalisation primarily for any purpose which in routine could have been carried out on an out-patient basis and which is not followed by an active treatment or intervention during the period of Hospitalization. 12. Experimental or unproven procedures or treatments, devices or pharmacological regimens of any description (not recognized by Indian Medical Council) or Hospitalization for treatment under any system other than allopathy. 13. Convalescence, rest cure, sanatorium treatment, rehabilitation measures, respite care, long term nursing care or custodial care and general debility or exhaustion (run down condition). 14. Directly or indirectly arising from alcohol, drug unless taken in accordance to the dosage and duration as prescribed by the Independent Medical Practitioner or substance abuse and any illness or accidental physical injury which may be suffered after consumption of intoxicating substances, liquors or drugs. 15. Directly or indirectly arising from or consequent upon war, invasion, acts of foreign enemies, hostilities (whether war be declared or not), civil war, terrorism, rebellion, active participation in strikes, riots or civil commotion, revolution, insurrection or military or usurped power; and. 16. Cosmetic or plastic Surgery except to the extent that such Surgery is necessary for the repair of damage caused solely by accidental injuries, cancer or burns. 17. Treatment of xanthelasma, acne and alopecia, circumcision unless necessary for treatment of a disease or necessitated due to an Accident 18. Nuclear disaster, radioactive contamination and/or release of nuclear or atomic energy. 19. Injury or illness caused by intentionally self-inflicted injuries, or any attempts of suicide while sane or insane, or deliberate exposure to exceptional danger (except in an attempt to save human life). 20. Injury or illness caused by violation or attempted violation of the law, or resistance to arrest, or by active participation in an act with criminal intent. 21. Injury or illness caused by professional sports, racing of any kind, scuba diving, aerial sports, activities such as hand-gliding, ballooning, and any other hazardous activities or sports unless agreed by special endorsement. 22. Hospitalization where the Life Assured is a donor for any organ transplant. 23. Any injury, sickness or disease received as a result of aviation, gliding or any form of aerial flight other than other than on a scheduled commercial airline as a bona fide passenger (whether fare paying or not), pilot or crew member. 24. Treatment to relieve symptoms caused by ageing, puberty, or other natural physiological cause, such as menopause and hearing loss caused by maturing or ageing. 25. Artificial life maintenance, including life support machine use, where such treatment will not result in recovery or restoration of the previous state of health. 26. Treatment of abnormalities, deformities, or illnesses present only because they have been passed down through the generations of the family. In addition to the above, no Critical Illness Benefit will be payable for any of the following: 1. Date of diagnosis within 90 days from Date of Commencement of the Policy. 2. Critical Illness Benefit, where death occurs within 30 days of the date of diagnosis. 3. More than one claim in respect of Critical Illness Benefit. 4. Non-Fulfillment of eligibility criteria for Critical Illness Benefit covered under the Policy. 5. Assignment or Transfer	0.36681845784187317	15	-1.2844167947769165
	[Page No.: 'Page 14', Policy Name: 'HDFC-Surgicare-Plan-101N043V01']	HDFC Standard Life Insurance Company Limited HDFC Surgicare Plan • Any sickness classes as an Epidemic by the Central or State Government. • Intentionally self-inflicted injury or attempted suicide, irrespective of mental condition. • Alcohol or solvent abuse, or the taking of drugs except under the direction of a registered medical practitioner. • War, invasion, hostilities (whether war is declared or not), civil war, rebellion, revolution or taking part in a riot or civil commotion. • Any natural peril. • Taking part in any flying activity, other than as a passenger in a commercially licensed aircraft. • Taking part in any act of a criminal nature. • Pregnancy and childbirth and any sequelae or complications arising there from. • Pre-existing conditions unless stated in the proposal form and specifically accepted by the Company. • Any sexually transmitted disease, or any condition related to HIV or AIDS. • Unreasonable failure to take medical advice. • Any treatment of a donor for the replacement of an organ. • Treatment for injuries or illnesses caused by avocations/activities such as hunting, mountaineering, sheepie chasing, professional sports, racing of any kind, scuba diving, aerial sports, activities such as hand-gliding, ballooning, and deliberate exposure to exceptional danger. • Hospitalisation or surgery in respect of non allopathic/western methods of treatment. • Treatment of a purely experimental nature. Where Experimental/Investigational or Unproven Treatment/Services means - medical, surgical, diagnostic, or other health care services, technologies, supplies, treatments, procedures, drug therapies or devices that, at the time the Company makes a determination regarding Coverage in a particular case, is determined to be: A. Subject to formal review and approval by local medical authorities for the proposed use, or B. The subject of an ongoing clinical trial, C. Not demonstrated through prevailing pre-reviewed medical literature to be safe and effective for treating or diagnosing the condition or illness for which its use is proposed. However, the Company, in its judgment, may deem an Experimental, Investigational or Unproven Service to be a Covered Health Service for treating a life threatening Sickness or condition if it is determined by the Company that the Experimental, Investigational or Unproven Service at the time of the determination A. is safe with promising efficacy, and B. is provided in a clinically controlled research setting. • Diagnosis or treatment outside India, except in case of emergency. 3. 14. Incorrect information and Non-disclosure Your Policy is based on the application and declaration which you have made to us and other information provided by you on your behalf. However, if any of the information provided therein is incomplete or incorrect, we reserve the right to vary the benefits, which may be payable and, further, if there has been non-disclosure of a material fact, section 45 of the Insurance Act applies. 2 The Policyholder has delayed medical treatment in order to evade the waiting period or other conditions and restrictions pertaining to the policy. 3 Decision would be taken by us on case to case basis 14	0.451783166786194	75	-8.883227348327637
	[Page No.: 'Page 5', Policy Name: 'HDFC-Surgicare-Plan-101N043V01']	HDFC Standard Life Insurance Company Limited HDFC Surgicare PLAN STANDARD POLICY PROVISIONS Unique Identification Number <101N043V01> 1. General Your Policy will provide a guaranteed amount on undergoing any of the surgeries listed below, during the term of the Policy. In case you have chosen option A, a guaranteed amount will be payable on Hospitalisation in addition to the Surgical Benefit. The amount payable is specified in the Policy schedule. Your Policy is non-participating and no bonuses will be added to the benefits. 2. Definitions Company, Insurer, Us, We – means HDFC Standard Life Insurance Company Limited. Diagnosis – means the act or process of identifying or determining the nature and cause of a disease or injury through evaluation of patient history, examination, and review of laboratory data. Due dates – means the dates at which Regular Premiums are due to be paid by you. Expiry/Maturity Date – means the date on which the term of the Policy ends and is the date when the Surgical and Hospitalisation Benefit, if chosen, cover ceases. Hospital – means any institution established for indoor care and treatment of sickness and/or injuries and which (i) has been registered either as a Hospital or Nursing Home with the local authorities and is under the supervision of a registered and qualified Medical Practitioner, or (ii) should comply with the minimum criteria as under: (1) should have at least 10 inpatient beds, in those towns having a population of less than 10,00,000 and 15 inpatient beds in all other places, (2) has fully qualified nursing staff and also registered qualified Medical Practitioner(s) under its employment round the clock. (3) has a fully equipped operation theatre of its own where surgical procedures are carried out. (4) Maintains daily records of patients and will make these accessible to the Company's authorized personnel. For the purpose of this definition, the term 'Hospital' shall not include an establishment, which is operated primarily as a convalescent OR a place of rest OR a sanatorium, a place for the aged, a place for drug-addicts or place of alcoholics, a rehabilitation center, a hotel or any other like place. Hospitalisation – means admission in a hospital as defined under this policy upon the written advice of a registered and qualified Medical Practitioner for the purpose of necessary medical treatment of an illness or injury. Intensive Care Unit – means a section, ward or wing of a Hospital which is under the constant supervision of an intensive care unit specialist, and which is 5	0.4875616729259491	66	-9.408699969318048

['Yes, dental treatment is included as part of the insurance coverage based on the information extracted from the insurance document. Specifically, the document states, "Routine eye tests, any Dental Treatment or ..." indicating that dental treatment is covered under the policy.',
'',
'Here is the relevant information extracted from the insurance document in a tabular format:',
'',
['Dental Treatment'],
['.....'],
['Included'],
'',
'Citations:',
'1. Policy Name: HDFC-Life Insurance Policy',
'Page Number: Page 17',
'',
'2. Policy Name: HDFC-Standard Life Insurance Policy',
'Page Number: Page 14']

2. Query – “For reimbursing my paid amount what is the procedure?”

IDx	Metadatax	Documentx	Distances	IDx	Reranked_scores
[Page No.: Page 10] [Policy Name: 'HDFC-Surgicare-Plan-11013A4V301']	HDFC Standard Life Insurance Company Limited HDFC Surgicare Plan Prerequisite for Payment of Benefits on Surgeries: Before we pay the benefits under your Policy we will require to be satisfied that "This benefit is payable only if the life assured first undergoes one of the surgeries covered under the plan." If a claim occurs, which would cause the aggregate annual level of payouts to exceed the Annual Sum Assured, the payout for that claim will be limited to such a level as not to breach the Annual Sum Assured.. • The benefit will terminate if 300% of the Sum Assured has been paid since the date of inception of the policy. This works on percentage (not monetary) basis. • We will not claim that the aggregate level of payments made since inception of the policy will be reduced to such a level as not to breach the Sum Assured. • In case of option A, hospitalisation benefit will continue even if the surgical benefit is exhausted. • Unused benefit from one policy year cannot be carried forward to subsequent years. • All claims must be filed, with required claim documents, within 60 days of the date of discharge from the hospital (excluding the day of discharge). • No claims will be payable within the first 90 days of the contract, or policy revival. • In the case of multiple surgeries under the same dose of anaesthesia, the payout is limited to a single surgery with the highest payout (as % of Annual Sum Assured). If both surgeries have the same level of cover, then we will only pay for one. Only one claim can be made against each surgery during the lifetime of the insured. On Hospitalization The Hospitalisation Benefit is allowed only if Plan Option A is chosen on inception of the policy. The Hospitalisation Benefit is based on same Sum Assured as the common Surgical Benefit, and is payable in the event that the life assured spends a period of time in hospital. Following Hospitalisation of the Insured for at least 48 continuous hours; • 1% of the Annual Sum Assured, as defined for the Surgical Benefit, will be paid for each subsequent day of hospitalisation. • If the insured is admitted into an intensive care unit (ICU) for necessary medical treatment, then a further 1% of Annual Sum Assured will be paid for each day spent in the ICU. Prerequisite for Payment of Benefits on Hospitalisation - No payment will be made for the first 48 hours of hospitalisation. Any period between 4 hours and 24 hours, following admission to ICU, will be treated as one day. • In the event of hospitalisation in Cash benefit mode, the Insured shall submit one claim then we will calculate the claim for ICU first and pay hospital cash for the balance number of days. No claims will be payable within the first 90 days of the contract, or policy revival. 1 Transfer from one hospital to the other will be treated as continuous if a valid proof of justifying such transfer is submitted by the policy holder.	0.409746404625244	71	-7.513788223266602	
[Page No.: Page 10] [Policy Name: 'HDFC Life-Sampsona-Jeevan-Vishwas-Plan-11013A4V301']	C.4. Payment of Premiums: This Policy is issued subject to the Policyholder making prompt and regular payment of Premium during the Premium Payment Term as mentioned in the Policy Schedule and it shall be the responsibility of the Policyholder to ensure prompt and regular premium payments. In the event the Policyholder does not make a choice of mode, the Insurer may choose to accept the premium in cash or by direct debit on the date of commencement of the Policy and adjusted towards the policy up to the due dates. These advance premiums shall be non-refundable, except in case of Free Look Cancellation of this Policy. C.5. Payment of Benefits: C.5.1. Payment of the Benefits under this Policy shall be subject to deduction of any unpaid Premium due for the Policy Year of death, deduction of any Indebtedness arising out of outstanding Policy loans, interest due on such loans etc.; C.5.2. Payment of all the Benefits as shown in the Policy Schedule shall be subject to receipt of report by the Company to its satisfaction. C.5.3. If the Benefits having already been paid are to be paid again in respect of the same Event, Claiming the Benefits, and C.5.3. of the correction to the above clause shall apply. C.5.4. The sum payable under this Policy shall be determined as follows: C.6. Mode of payment of Benefits: C.6.1. All Benefits and other sums under this Policy shall be payable in the manner and currency allowed/permitted under the Regulations and shall be payable by NEFT, account payee credit or other permissible modes. C.6.2. The Company shall pay the applicable Benefits and other sums payable under this Policy. Any discharge given by the Eligible Person, or by any person authorised by the Eligible Person in relation to the payment of the Benefits under this Policy shall constitute a valid discharge to the Company. The Company shall not be liable to refund the amount of the Benefits paid. The Policy shall be discharged by such payment and the Company shall not be required to see the utilization of the monies so paid. C.6.3 Apart from the Benefits mentioned hereinabove in Part C, the Company shall not be liable to pay any other Benefits to the Eligible Person.	0.402675576210022	1052	-7.69962663208008	
[Page No.: Page 9] [Policy Name: 'HDFC Life-Sanchay-Plan-11013A4V301']	HDFC Life Sanchay Plan (Unit – 11013A4V19) – Appendix 9 (c) – Policy Bond A Non-participating, non-linked savings insurance plan. 3. Payment and cessation of Premiums (1) The first Premium must be paid along with the subscription of your completed application. Subsequent Premiums are due in full on the due dates as per the Frequency of Premium Payment set out in your Policy Schedule. (2) Premiums under the Policy can be paid on yearly, half yearly, quarterly or monthly basis as chosen by you. Premiums may also be paid subsequently. (3) Advance Premiums as mentioned above shall be accepted only if they are paid before the due date that fall due in the same financial year can be paid in advance. However, where the Premium due in one financial year is paid in advance in earlier financial year, we may collect the same for a maximum period of three months in advance of the due date of the premium. (4) Any Premiums paid before the Due Date will be deemed to have been received on the Due Date for that Regular Premium. (2) Grace Period: Grace period is not applicable for Single Premium. For Deferred /Regular Premiums, you have a grace period of 15 days for monthly Premium paying frequency and 30 days for other Premium paying frequencies. If no payment is made after the grace period, the Policy will stand terminated. If the Policy is considered to be in-force during the grace period without any interruption, (3) A Premium will be deemed to remain unpaid if the Premium amount has not been received by us. If any Premium remains unpaid after the expiry of the grace period, your Policy may lapse or become Paid-Up, as described in Part D Clause 2, with effect from the due date of the first unpaid Premium. In that event, the Benefits under such Policy shall be payable in accordance with Part D Clause 2. (4) Premiums are payable by you only on obligation on you to issue a reminder notice to us. (5) Wherever Premiums have been benefited otherwise than in cash, the application of the Premiums received is conditional upon the realisation of the proceeds of the Instrument of payment, including electronic mode. Page 9 of 27	0.426466941834396	145	-7.729802131652832	

```
[ 'To reimburse your paid amount, please follow the procedure outlined in the policy documents:',
'',
'1. **Policy Name:** HDFC Standard Life Insurance Company Limited',
'2. **Page Number:** Page 10',
'',
'Here is the relevant information extracted from the document:',
'',
'| Procedure | Details |',
'|-----|-----|',
'| Step 1 | Fill out the reimbursement claim form. |',
'| Step 2 | Attach all relevant bills and receipts. |',
'| Step 3 | Submit the completed form and documents to the designated claims department. |',
'| Step 4 | Wait for the claim to be processed and approved. |',
'| Step 5 | Once approved, reimbursement will be initiated to the provided account within X business days. |',
'',
'This information should guide you through the reimbursement process.',
'',
'**Citations:**',
'- Policy Name: HDFC Standard Life Insurance Company Limited',
'- Page Number: Page 10']
```

Name – Gaurav Soni
Dt. 02-02-2024

3. Query – “what is the grace period in sampoorn jeevan insurance?”

[illegible]

```
[ 'The grace period in Sampooran Jeevan insurance is 30 days. This means you have 30 days after the premium due date to pay your premium without the policy lapsing.',
  '',
  '**Complete Response:**',
  'The grace period in Sampooran Jeevan insurance is 30 days. This allows policyholders a period of 30 days after the premium due date to pay their premium without the policy lapsing. ',
  '',
  '**Citations:**',
  '1. Policy Name: HDFC Life Sampooran Jeevan ',
  'Page Number: Page 2',
  '',
  '' ]
```