

# 교육 과정 소개서.

---

고성능 RAG를 위한 66가지 최신 RAG 테크닉 (ft.  
Advanced RAG, Modular RAG)



## 강의정보

강의장	온라인 강의   데스크탑, 노트북, 모바일 등
수강 기간	평생 소장
상세페이지	<a href="https://fastcampus.co.kr/data_online_ragupskill">https://fastcampus.co.kr/data_online_ragupskill</a>
강의시간	29시간 7분
문의	<a href="#">고객센터</a>

## 강의특징

나만의 속도로	낮이나 새벽이나 <b>내가 원하는 시간대</b> 에 나의 스케줄대로 수강
------------	---

원하는 곳 어디서나	시간을 쪼개 먼 거리를 오가며 오프라인 강의장을 찾을 필요 없이 <b>어디서든 수강</b>
---------------	---

무제한 복습	무엇이든 반복적으로 학습해야 내것이 되기에 이해가 안가는 구간 <b>몇번이고 재생</b>
-----------	--



## 강의목표

- RAG 파이프라인의 단계별 성능 평가, 그리고 그 결과를 기반으로 한 단계별 성능 최적화 전략 수립 방법을 학습합니다.
- 현직 신한AI, SK디스커버리에 근무 중인 AI 전문가들이 실무에서 직접 연구 및 사용 중인 최신 RAG 테크닉을 알려드립니다.
- RAG를 구축해본 경험이 있으신 분들, 특히 Naive RAG를 구축하며 그 한계를 실감하고 계신 현직자 분들께 추천드립니다.
- 오직 패스트캠퍼스에서만 만나볼 수 있는 2024년 기준 최신 RAG 테크닉 강의입니다. RAG의 3가지 대표적 패러다임은 물론, 학계에서 막 떠오른 최신 테크닉까지 만나보세요.

## 강의요약

- 최신 논문을 분석하며 Advanced RAG, Modular RAG 포함 66가지 RAG 테크닉과 이를 조합하는 방법을 배웁니다.
- 2024년 학계에서 화제가 된 논문 6가지를 분석하며, 기존 RAG 테크닉의 단점을 보완한 신생 테크닉에 대해 알아봅니다. 비교적 유명한 SELF-RAG, CRAG부터 다른 어떤 강의에서도 다루지 않는 Multi-hop RAG, RAPTOR-RAG 등 최신 개념까지 만나보세요.
- RAG 파이프라인을 최적화하기 위해서는, 파이프라인의 각 단계에서 어떠한 모듈이 성능 저하를 일으키는지 성능 평가를 통해 파악하고 이에 걸맞는 최적화 전략을 수립하여야 합니다. RAG 성능 평가 자동화부터 그 결과에 따른 단계별 최적화 전략을 깊이있게 알아봅니다.



## 강사

정민수

과목

- 고성능 RAG를 위한 66가지 최신 RAG 테크닉 (ft. Advanced RAG, Modular RAG)

약력

- 전) 신한은행 NLP 개발
- 전) 크래프트테크놀로지스 NLP Researcher

곽기은

과목

- 고성능 RAG를 위한 66가지 최신 RAG 테크닉 (ft. Advanced RAG, Modular RAG)

약력

- 현) SK디스커버리 Data Scientist
- 전) MunichRe Data Scientist

서지영

과목

- 고성능 RAG를 위한 66가지 최신 RAG 테크닉 (ft. Advanced RAG, Modular RAG)

약력

- 현) Microsoft Data&AI Specialist
- 전) 한국은행 Data

## CURRICULUM

## 01.

RAG 기초 이론  
& 실습

파트별 수강시간 03:57:59

<b>CH01. 기업은 왜 챗GPT 서비스를 사용할 수 없나?</b>
01. 챗GPT 서비스의 제약
02. 기업에서 챗GPT를 안전하게 사용할 수 있는 방법은?
<b>CH02. RAG 개요</b>
01. 파인튜닝 개요
02. RAG 개요
03. RAG는 언제 사용하는가?
04. RAG 구현시 데이터 흐름
<b>CH03. 청킹(Chunking)</b>
01. 청킹 개념
02. 청킹 유형 및 고려사항
<b>CH04. 검색 유형 및 방법</b>
01. 검색 유형
02. 검색 방법
<b>CH05. 벡터 데이터베이스</b>
01. 벡터 데이터베이스 개념
02. 벡터 데이터베이스 유형
03. 벡터 데이터베이스 사용 방법
<b>CH06. LLM에서 사용하는 벡터 검색</b>
01. 벡터 데이터베이스 검색
02. 벡터 데이터베이스 인덱스
03. Reranking
<b>CH07. RAG 구현시 고려사항</b>
01. 답변 속도, 정확도 및 모니터링
<b>CH08. RAG 간단한 실습</b>
01. RAG 기반의 챗봇 만들기
02. RAG 기반의 챗봇 애플리케이션 만들기

## CURRICULUM

## 02.

# Naive RAG로 인재 추천 시스템 만들기

파트별 수강시간 03:37:11

<b>CH01. VectorDB Overview</b>
01. 개요
02. Indexing
03. Querying
04. 사용 가능 Vector DB 리스트
05. Pinecone을 활용한 Indexing & Searching
<b>CH02. LLM Architecture</b>
01. LLM architecture, Cheat sheet 소개
<b>CH03. Pinecone을 활용한 인재 추천 시스템 만들기</b>
01. Search 및 사용자 인풋 정의



CURRICULUM

03.

고성능 RAG를  
위한 최신 테크닉

파트별 수강시간 03:37:13

CH01. 강의 개요
01. 강의 개요
CH02. Modular RAG
01. 효율적인 RAG 아키텍처 - modular RAG
CH03. 다양한 RAG 테크닉
01. 정보를 re-group하는 Raptor rag (이론)
02. 정보를 re-group하는 Raptor rag (실습)
03. reflection token을 사용하는 self rag (이론)
04. 정보를 한 번 더 검토하는 Corrective rag (이론)
05. 정보를 한 번 더 검토하는 Corrective rag (실습)
06. 질문의 수준에 따라 , adaptive rag (이론)
07. multi doc retrieval benchmark, multi-hop rag (이론)
08. 정보의 구조화, Tree rag (이론)
09. Modular rag의 다양한 테크닉 (1)
10. Modular rag의 다양한 테크닉 (2)

## CURRICULUM

## 04.

# Advanced RAG로 실시간 뉴스 정리봇 & 투자 어시스턴트 만들기

파트별 수강시간 07:44:17

<b>CH01. Intro</b>
01. 강의 개요
02. 프로젝트 아키텍처 설명
03. 프로젝트 환경 설정
<b>CH02. Named Entity Recognition</b>
01. 엔티티 정의
02. 분류 기반의 개체명 인식기 (spaCy)
03. 생성 기반의 개체명 인식기 (LLM)
<b>CH03. Entity Linking</b>
01. 엔티티 링킹 (Entity Linking)
02. 다양한 기업 표현 확인하기
03. 엔티티 링킹 구현하기
<b>CH04. Aspect-based Sentiment Analysis</b>
01. 감성분석 소개
02. 분류 기반의 감성 분석
03. 생성 기반의 감성 분석
04. 감성 분석 soft-labeling
05. 엔티티 중심의 감성 분석 구현하기
06. 엔티티 링킹과 감성 분석 연결하기
<b>CH05. 실시간 파이프라인 구축</b>
01. 아키텍처 설명
02. 실시간 뉴스 수집
03. MongoDB 데이터베이스 연결하기
04. 분석 파이프라인 만들기
05. 데이터베이스에 결과 적재하기
<b>CH06. Streamlit을 사용한 대쉬보드 제작</b>
01. 시각화 프로젝트 소개
02. Streamlit 기능 소개
03. Streamlit 대쉬보드 제작하기





CURRICULUM

04.

**Advanced  
RAG로 실시간  
뉴스 정리봇 &  
투자 어시스턴트  
만들기**

파트별 수강시간 07:44:17

<b>CH07. 요약&amp;이벤트 추출 Intro</b>
01. 강의 개요
02. 프로젝트 아키텍처 설명
<b>CH08. 뉴스 카테고리 분리하기</b>
01. 뉴스 카테고리 분류 체계
02. 뉴스 API 를 이용한 실시간 뉴스 수집
03. 뉴스 카테고리 분류하기
<b>CH09. 기업 공시 정보 분석하기</b>
01. 미국 증권거래위원회 (SEC) 소개
02. SEC 공시 정보 파싱하기
03. 공시텍스트 분류하기
04. 한국 전자공시시스템 DART 소개
<b>CH010. 프롬프트엔지니어링을 활용한 요약 및 이벤트 추출</b>
01. 아키텍처 설명
02. 뉴스 및 공시 요약 프롬프트 엔지니어링
03. 이벤트 추출 프롬프트 엔지니어링
04. Reasoning step을 적용한 프롬프트 엔지니어링
05. Few-shot Learning을 이용한 프롬프트 엔지니어링
<b>CH011. PoC 제품화하기</b>
01. 아키텍처 설명
02. 분석 파이프라인 만들기
03. 데이터베이스에 결과 적재하기
04. Streamlit 시각화 페이지 설계
05. Streamlit 대쉬보드 제작하기

## CURRICULUM

## 05.

# Modular RAG로 인사 시스템 만들기

파트별 수강시간 05:08:57

## CH01. Intro

01. 강의 개요
02. 프로젝트 아키텍처 설명

## CH02. 챗봇 프롬프트 엔지니어링

01. 챗봇 페르소나 설명
02. 페르소나 프롬프트 작성
03. 혐오표현 방지 프롬프트 작성
04. 보안 프롬프트 작성

## CH03. RAG 아키텍처 소개

01. 아키텍처 설명
02. LangChain 프레임워크 소개

## CH04. 벡터 검색 구현하기

01. 임베더 개념 설명
02. API 기반 문서 임베딩
03. PyPDF 기반 PDF 문서 파싱
04. LangChain Splitter 문서 쪼개기
05. LangChain Retriever 검색 구현
06. 단어 기반의 검색 알고리즘 설명
07. BM25를 결합한 하이브리드 서치 구현
08. 검색 성능 평가

## CH05. RAG 아키텍처 구현하기

01. 아키텍처 설명
02. RAG 프롬프트 엔지니어링
03. RAG 코드 구현하기
04. LLM 분석루프 추가하기
05. 비정형정형 데이터 기반의 RAG 구현하기

## CH06. 온프레미스 RAG 구현하기

01. 아키텍처 설명
02. SentenceTransformers 사용하기
03. SBERT 모델 선정하기
04. SBERT 임베딩 로직 구현하기
05. Huggingface sLLM 사용하기
06. Huggingface Transformers 라이브러리 설명
07. sLLM을 활용한 문장 생성
08. sLLM을 활용한 RAG 아키텍처 구현
09. 실시간 온프레미스 환경 구축하기
10. Streamlit 채팅 데모 구현하기

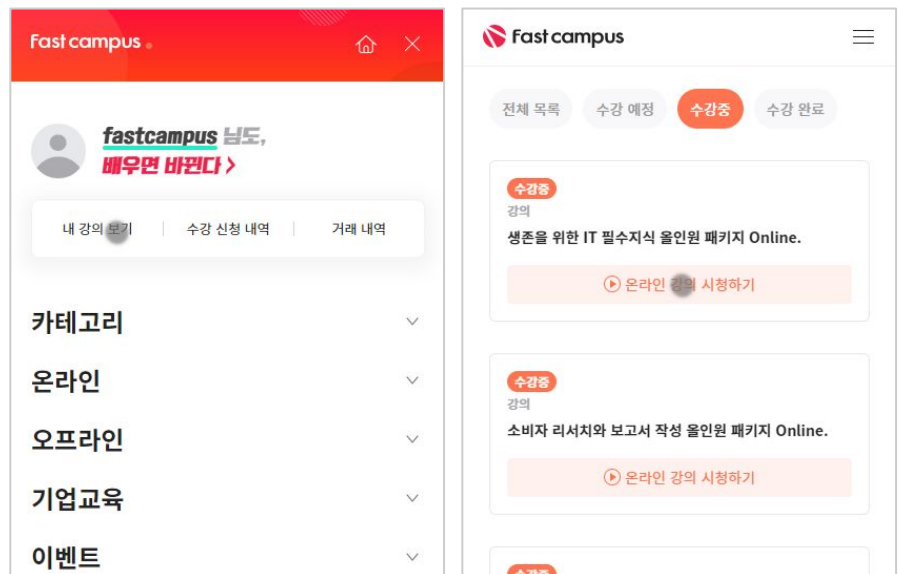


## 주의 사항

- 상황에 따라 사전 공지 없이 할인이 조기 마감되거나 연장될 수 있습니다.
- 패스트캠퍼스의 모든 온라인 강의는 아이디 공유를 금지하고 있으며 1개의 아이디로 여러 명이 수강하실 수 없습니다.
- 별도의 주의사항은 각 강의 상세페이지에서 확인하실 수 있습니다.

## 수강 방법

- 패스트캠퍼스는 크롬 브라우저에 최적화 되어있습니다.
- 사전 예약 판매 중인 강의의 경우 1차 공개일정에 맞춰 '온라인 강의 시청하기'가 활성화됩니다.



## 환불 규정

- 온라인 강의는 각 과정 별 '정상 수강기간(유료수강기간)'과 정상 수강기간 이후의 '복습 수강기간(무료수강기간)'으로 구성됩니다.
- 환불금액은 실제 결제금액을 기준으로 계산됩니다.

수강 시작 후 7일 이내	100% 환불 가능 (단, 수강하셨다면 수강 분량만큼 차감)
수강 시작 후 7일 경과	정상(유료) 수강기간 대비 잔여일에 대해 환불규정에 따라 환불 가능

※ 강의별 환불규정이 상이할 수 있으므로 각 강의 상세페이지를 확인해 주세요.