



제 13회 투빅스 컨퍼런스

ToBiz

# 구직자를 위한 비즈니스 길잡이

조준혁 이수민 김동현 김송민 김주호 정수연



# CONTENTS INDEX

---

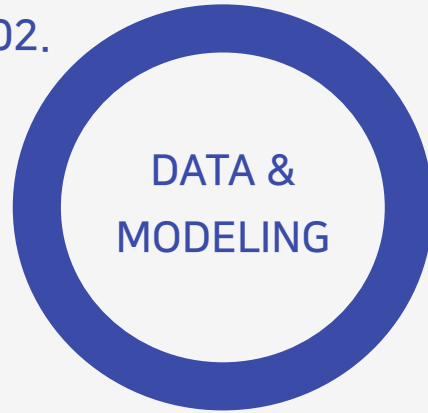
01.



INTRO

주제 선정 배경  
문제 정의 & 해결 방안

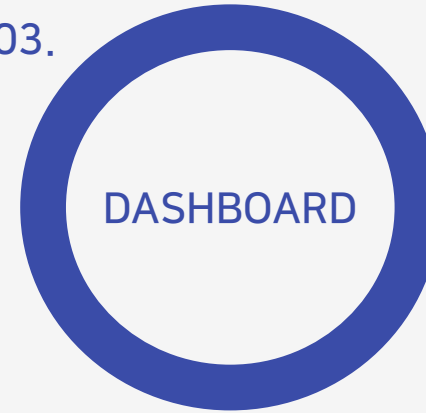
02.



DATA &  
MODELING

데이터 수집  
전처리 & 모델링

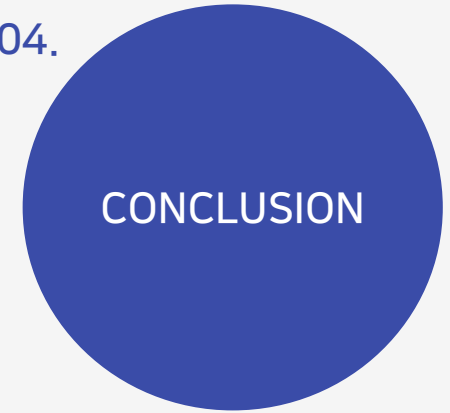
03.



DASHBOARD

대시보드 설명 & 시연

04.



CONCLUSION

기대효과



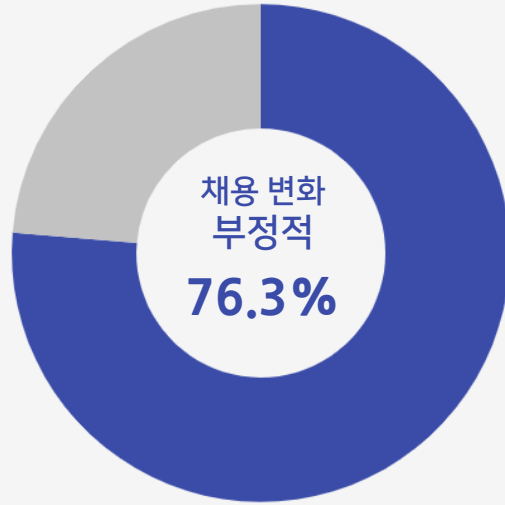
# INTRO

주제 선정 배경 | 문제 정의 & 해결 방안



### 최근 코로나로 인해 얼어붙은 취업시장

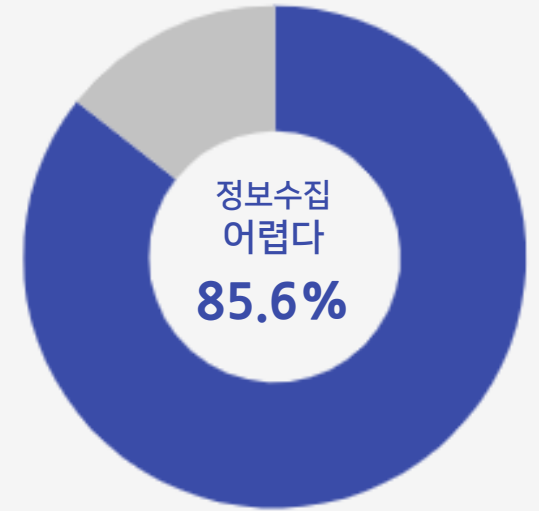
구직자 10명 중 9명이 코로나로  
인한 채용시장 변화 체감  
그 중 부정적 변화에 대한 체감이  
75% 이상 (사람인, 2021)



### 취업 관련 정보 부족

구직자 85.6%,  
“기업 관련 정보 수집이 어렵다”  
(잡코리아, 2021)

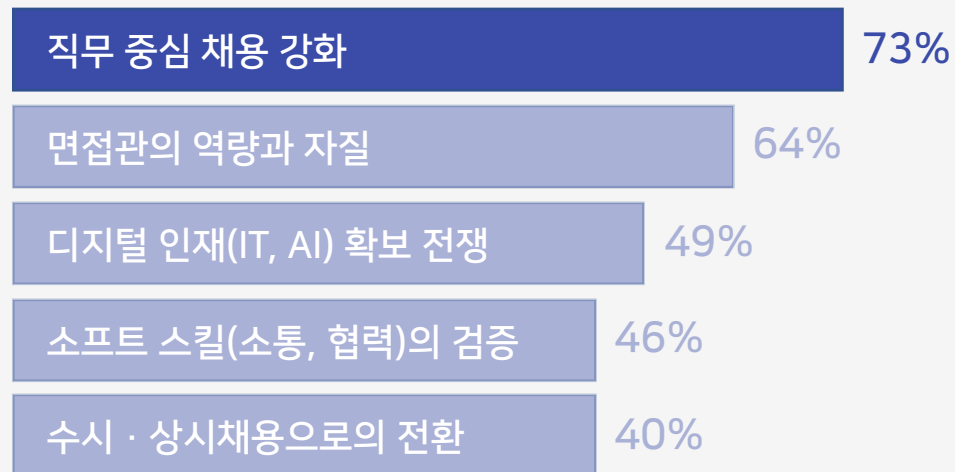
언택트 채용에 대한 부담요소의  
45.3%가 관련 정보 부족 때문  
(사람인, 2020)



## 직무역량 중심 채용

직렬정도만 정하고 취업을 준비하던 과거와 달리, 직무역량 중심 채용 대두되면서 직무 세분화되고 중요성 더욱 증대  
→ 구직자들이 직무를 먼저 선택하고 취업을 준비하는 경향이 강화

### 면접관이 꼽은 채용트렌드 2022



⋮

### 수시채용의 핵심 '직무적합성'...이렇게 준비하면 합격

▲ 이상진 기자 | ⓒ 입력 2021.08.30 16:48 | ⓒ 수정 2021.08.30 16:59 | 댓글 2

#### [초점] 기업 채용시 '직무능력' 가장 높게 평가해..."

ⓒ 김민서 기자 | ⓒ 승인 2022.01.12 08:05 | 댓글 0

신입 채용 시 전공의 직무 관련성이 학력보다 중요  
경력직 채용 시 직무 관련 업무 경험 가장 중요해  
정부, 개인 직무능력 정보 한눈에 볼 수 있는 '직무능력은행제' 구축  
일경험 프로그램 참여 등으로 직무능력 함양 시켜야 성공취업 가능해져



### 구직자를 위한 길잡이가 필요해!

직무 중심 채용 트렌드에 맞춰, 희망 진로와 직무를 찾을 수 있도록 돕는  
진로 상담 서비스 등은 워크넷 등의 플랫폼을 통해 활발하게 이루어지고 있는 편이다.

그러나 직무를 정한 이후에도 스스로 정보를 탐색해 나가야 하며,  
구직자들은 정보 부족 문제로 어려움을 겪고 있다.

직무를 정한 구직자가 채용까지의 과정에서 필요로 하는 정보들을  
하나의 대시보드 흐름으로 제공할 수는 없을까?"



## 문제 정의 & 해결방안

01.

내 직무의 채용공고 모아 보고  
최근 채용 트렌드 알아보기

내가 희망하는 직무의 채용 공고들을 한 눈에 확인하기 어렵다. 채용 트렌드 중 새롭게 나타난 내용이 있는지 알기 어렵다.



최근 몇년간 있었던 채용공고와 채용 트렌드 제시! 채용 시장을 빠르게 파악하고 집중적으로 어필할 부분을 파악하여 효율적인 구직 준비를 시작할 수 있다.

02.

자소서 기반으로  
합격 확률 높은 기업 & 포지션 추천 받기

어느 자리가 나에게 잘 맞을 지 알 수 없어서 이력서를 난사하게 된다. 이 회사, 이 직무에 붙을 수 있을 것인지 예측이 어렵다.



자소서 기반으로 합격 확률 높은 기업과 직무 추천! 어디에 주력해야 할지, 틈새 시장이 어디인지 알고 해당 회사들을 집중 공략할 수 있다.

03.

추천 받은 기업들 + 다양한 기업들의  
정보를 모아 보기

연봉, 워라밸 등 기업 정보와 리뷰를 보기 위해 잡플래닛, 블라인드, 캐치 등 여러 플랫폼을 살펴보느라 리서칭이 한참 걸린다.



분산되어 있는 기업 정보 한 눈에 알아볼 수 있도록 여러 플랫폼의 기업 정보와 리뷰를 통합! 리서칭에 뺏기는 에너지를 절약하고 구직 준비에 집중할 수 있다.



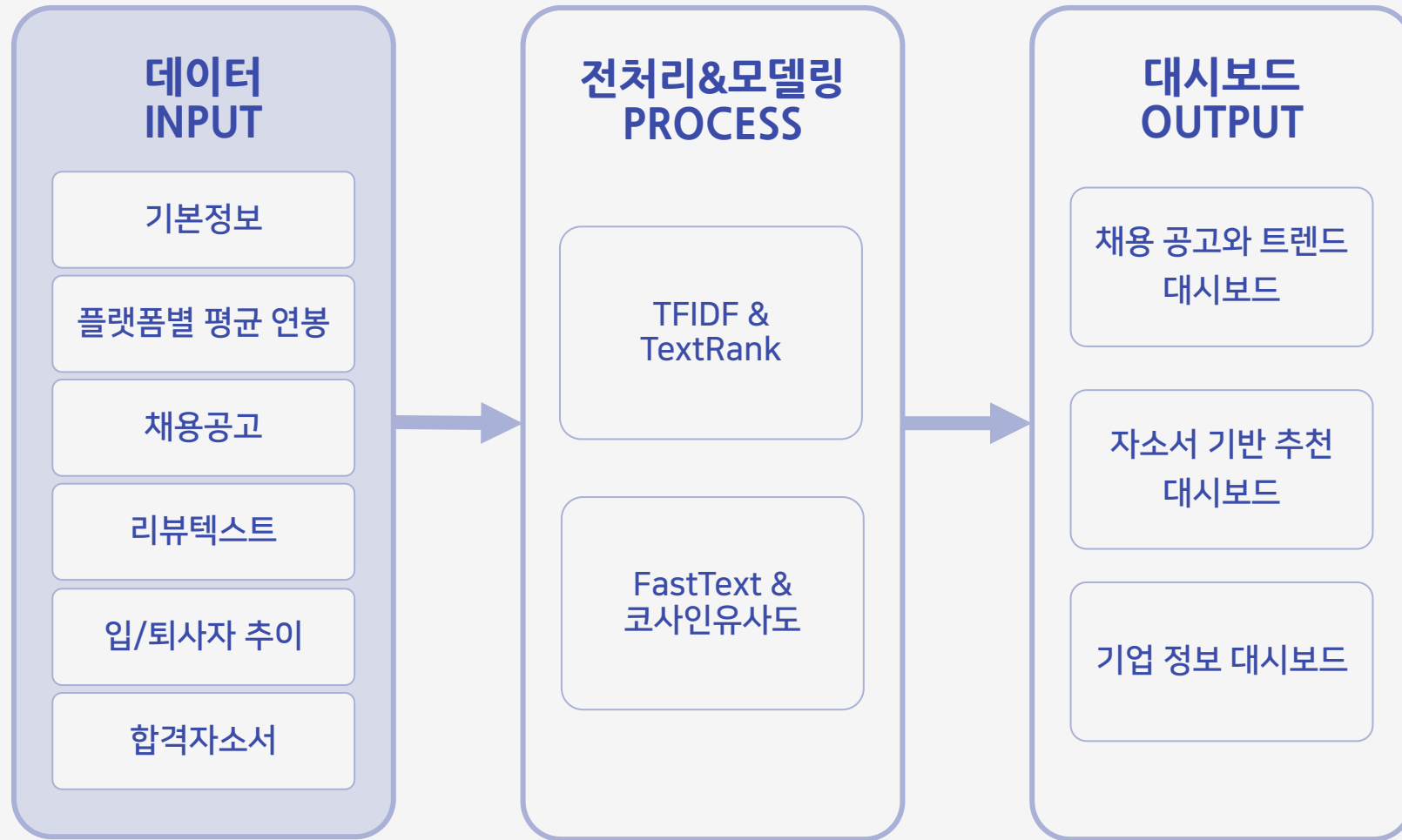
# Data & Modeling

데이터 수집 | 전처리 & 모델링





## WHERE WE AT?



# 10개 업종 1000개 기업 선정

### 잡플래닛 분류 기준에 따라 10개 업종 선택

같은 직무도 업종별로 특성 달라서 업종 정보 포함할 필요성  
가장 이용자와 정보량이 많아서 대표성이 있는 '잡플래닛' 플랫폼의 업종 분류를 따라 전체 업종을 10 분류로 나눔

중공업 계열의 품질관리 직무와 반도체 계열의 품질관리 직무는 완전히 다르거든요. (중략) 이런 까닭에 업종의 선택을 가장 명확하게 하는 게 우선돼야 한다고 당부하고 싶습니다.

이대성 대통령 정책기획위원회 자문위원 인터뷰 中

### 잡플래닛 리뷰 많은 순으로 각 업종 당 100개 기업 선정

구직자가 가장 궁금한 것은 현직자의 솔직한 이야기  
신뢰도를 확보하기 위해 리뷰수가 많은 순으로 100개 기업 선정

기업에서 근무하고 있는 직원을 초청해 해당 기업에 대한 궁금증을 풀어주는 현직자 멘토링은 3661명(전체의 37.0%)의 구직자가 신청했다. 현직자 멘토링은 이전부터 구직자들에게 인기 높은 취업 준비 프로그램이다.

진학사 캐치, 2021



## 7개의 다양한 관련 플랫폼 선정

**blind** TOPIC

**LINKareer** 링커리어

끝이 다른 시작  
**JOBKOREA**

**Jobplanet**



**KREDIT JOB**

**saramin**

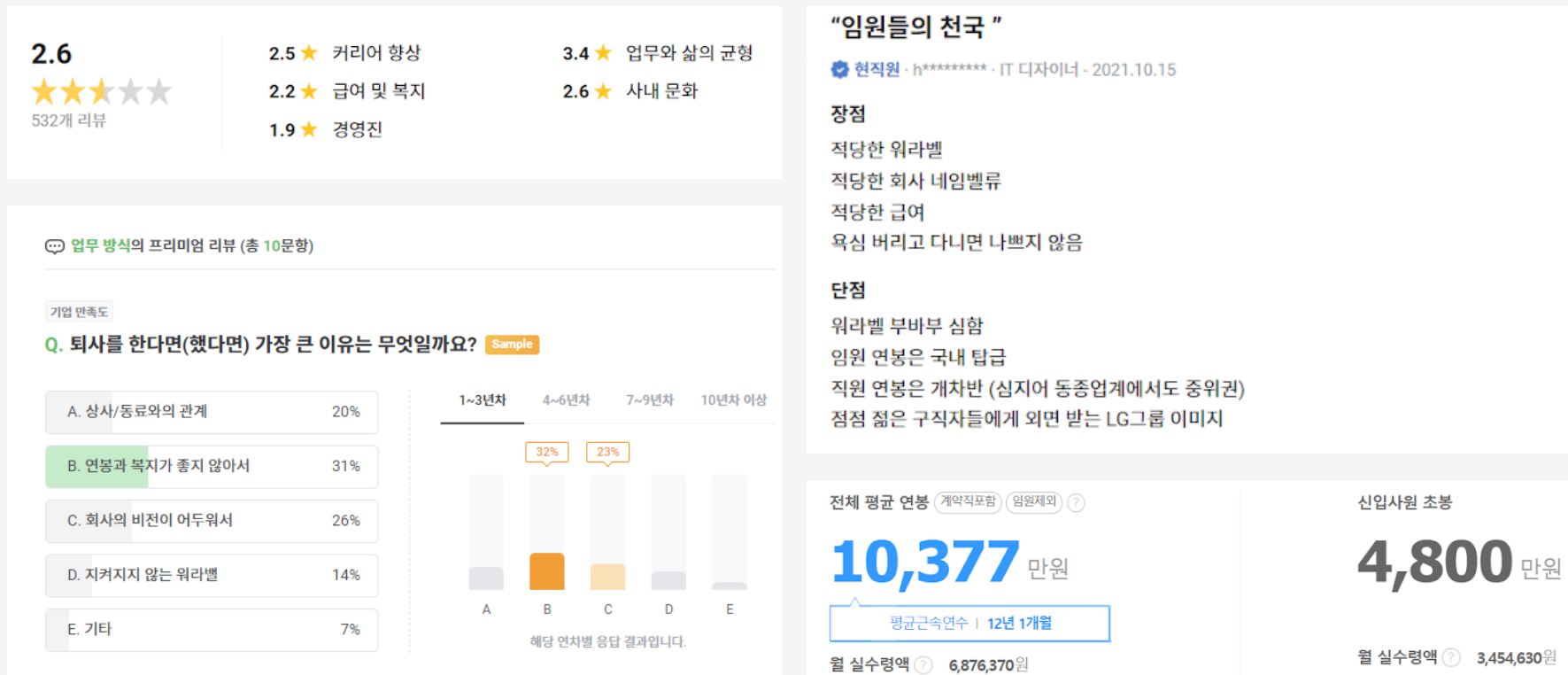
**CATCH**



## 기본정보, 플랫폼별 평균 연봉, 입/퇴사자 추이, 기업 로고 이미지, 리뷰 평점, 리뷰 텍스트

캐치, 블라인드, 잡코리아, 크레딧잡, 사람인

특히 잡플래닛, 캐치, 블라인드에서 현직자 리뷰 데이터 338,138개 수집



## 직무 대분류별 합격자소서 링크리어, 잡코리아

보다 대표성 있는 플랫폼인 링크리어의 직무 대분류를 기준으로 직무 대분류별 11,195개 합격자소서 수집  
선정한 1000개 기업에 해당하는 자소서들만 추출

## 채용 공고 잡플래닛

선정한 1000개 기업들에서 낸 36,482개의 채용 공고 수집

검색결과 788건 관련도순 ▾

합격연도	기업명	직무명	합격스택/내용	스크랩
2021 상반기	KB부동산선택	경영	한국외대 / 경영학전공 / 학점 3.43 / 토익: 865, 토익스피킹: 140/레벨 6, 오픽: , 기타: / 사회생활 경험: / ... 1. 본인이 지원한 직무에 적합한 이유에 대하여 서술해 주시기 바랍니다. (500자 이내) 남자만 삼 형제 집안에서 둘째로 태어나 먹을 것부터 시작에 사소한 모든 것에서 싸움이 끊이지 않았습니다. 어느 날은 동생으로서, 어느 날은 형으로서 부모님의 꾸중을 듣고 혼이 났었던 유년기가 기억이 납니다. 부모님께서 그때마다 `...	1
2021 상반기	국토안전관리원	경영	부경대 / 산업공학과 / 학점 3.26 / 토익: 790, 토익스피킹: , 오픽: , 기타: / 사회생활 경험: 공기업 인턴 / ... 1. 최근 국토안전관리원의 주요 이슈에 대해 한 가지를 선택하여 그것에 대한 본인의 의견을 기술하여 주십시오. 국토안전관리원은 건설공사 및 시설물의 안전 유지와 체계를 확립하고, 시설관리자들의 안전관리 이행력을 높이는 책임을 가진 기관입니다. 특히 최근엔 건설사고 사례 분석을 통한 '위험요소 프로파일'을 작성함...	2
2020 하반기	서울주택도시...	사무직_경영...	국민대 / 영어영문학부 / 학점 3.99/4.5 / 토익: 945 / 사회생활 경험: 국민건강보험공단 인턴 3개월, 건강... 서울주택도시공사에 지원하게 된 동기는 무엇인지 기술해 주십시오. * [500자 이내] ① 다른 기관이 아닌 서울주택도시공사/해당직무를 지원한 이유는 무엇입니까? ② 귀하의 감정이나 역량 중에서 어떤 것이 향후 서울주택도시공사 발전에 기여할 수 있다고 생각하십니까? 487/500 (글자 수, 공백 포함) 제 직업관은 누군가...	15
계명대학교 / 정치외교학 / 학점 4.11/4.5 / 토익: 905, 오픽: IH / 사회생활 경험: 인턴 1년 8개월 / 한국사...				

링크리어 합격자소서 검색 결과

kakao [잡플래닛 매칭] 추천 시스템 연구/개발  
(주)카카오

### 주요업무

#### ◆ 조직소개

- 추천기술파트는 카카오의 여러 서비스에 필요한 추천시스템과 머신러닝 기술을 활용한 분류/학습기를 제공하고 하고 있음
- State of the art 수준의 기술을 실제 서비스에 적용할 기회가 열려있음
- 하나의 서비스 기능 개발하기 위한 거의 모든 기술적 스택을 담당함
- 스타트업과 같은 자유로움과 분위기로 일하면서 포근한 카카오의 크루가 될 수 있음

#### ◆ 업무내용

- 텍스트/음원/영상등의 콘텐츠와 사용자 활동 정보를 활용한 추천시스템 알고리즘 개발 및 분석
- 대용량 처리를 위한 알고리즘 개발
- 서비스에서 겪는/요구하는 다양한 문제에 대한 해결책과 솔루션 제공
- 최신 추천 시스템 연구 분석과 서비스화 가능성에 대한 실험

### 자격요건

- 적어도 하나 이상의 능숙한 프로그래밍 언어가 있으신 분

잡플래닛 채용 공고 예시



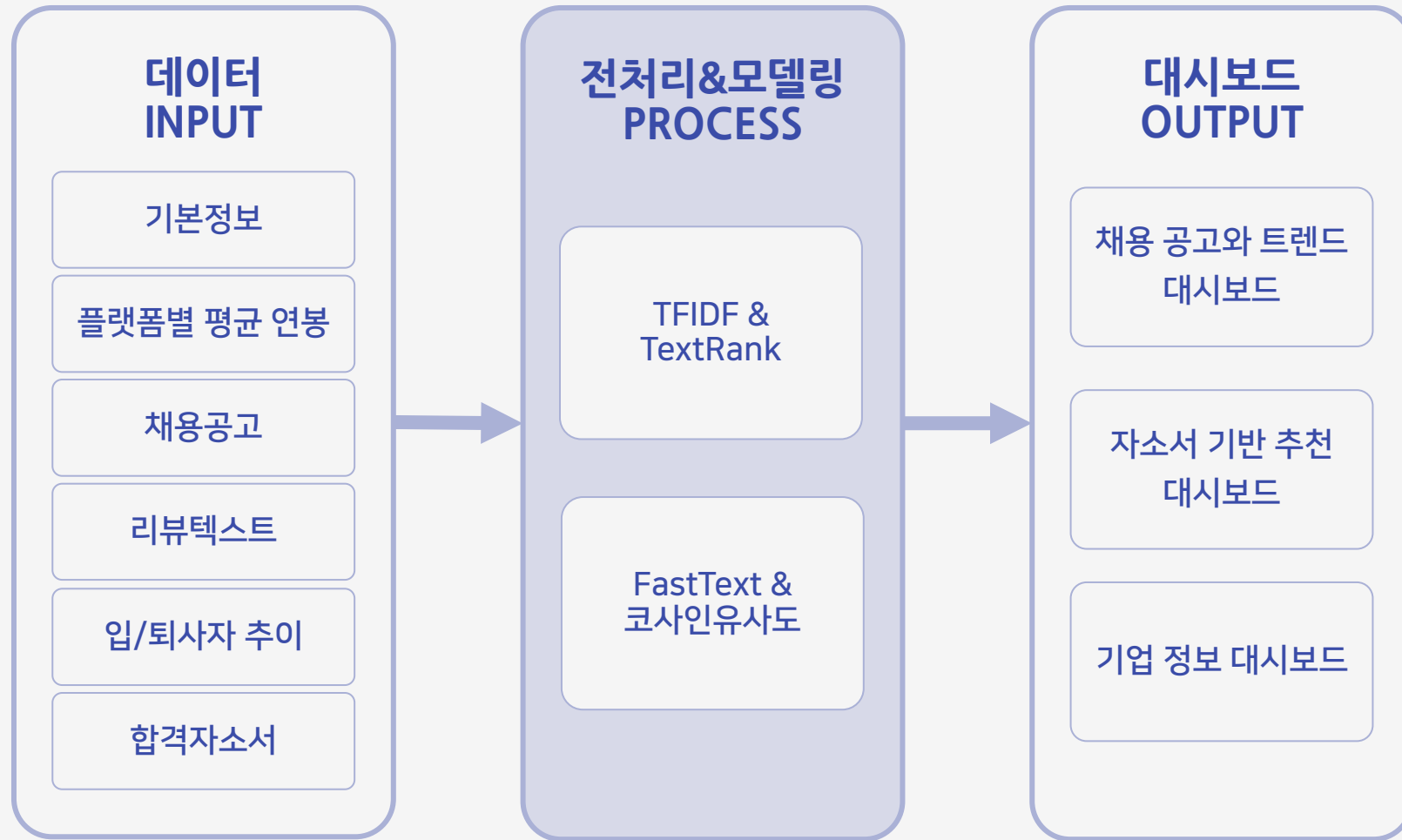
## => 스키마 작성, 형태에 맞게 크롤링

	A	B	C	D	E	F	G	H	I	J	K	L	M	
1	기본 정보데이터													
2	회사명	연봉	워라밸	커리어	조직문화									
3														
4		산업	기업형태	직원수	설립	대표	매출	주소	웹사이트	연혁	소개			
5														
6		로고URL												
7														
8		초봉												0번
9														0번
10		매우불만족	불만족	보통	만족	매우만족								1번
11														2번
12		입사율	퇴사율	평균근속연수										3번
13														4번(초봉)
14		부채비율	당좌비율	신용등급										4번(플랫폼별 평균 연봉)
15														5번
16		기업 신용등급	(fn가이드)유동비	(fn가이드)당좌비	(fn가이드)부채비	(fn가이드)유보율	(fn가이드)순차율	(fn가이드)이자보	(fn가이드)자기자본비율					5번
17														6번(고용안정성)
18	플랫폼별 평균 연봉 데이터				월별 총인원, 입, 퇴사자 추이									6번(고용안정성)
19	회사명	플랫폼	평균연봉		회사명	월	총인원	입사자수	퇴사자수					6번(재무안정성)
20														6번(재무안정성)
21	ex)				ex)									
22	NC소프트	캐치	8,728		NC소프트	2021-02-01	4,431	158	76					
23	NC소프트	잡코리아	10,550		NC소프트	2021-03-01	4,424	69	39					
24	NC소프트	사람인	10,547		NC소프트	2021-04-01	4,478	93	45					
25	NC소프트	크레딧잡	6,556		NC소프트	2021-05-01	4,538	105	60					
26														

스키마 중 일부



## WHERE WE AT?



## 형태소 분석 & 키워드 추출 프로세스





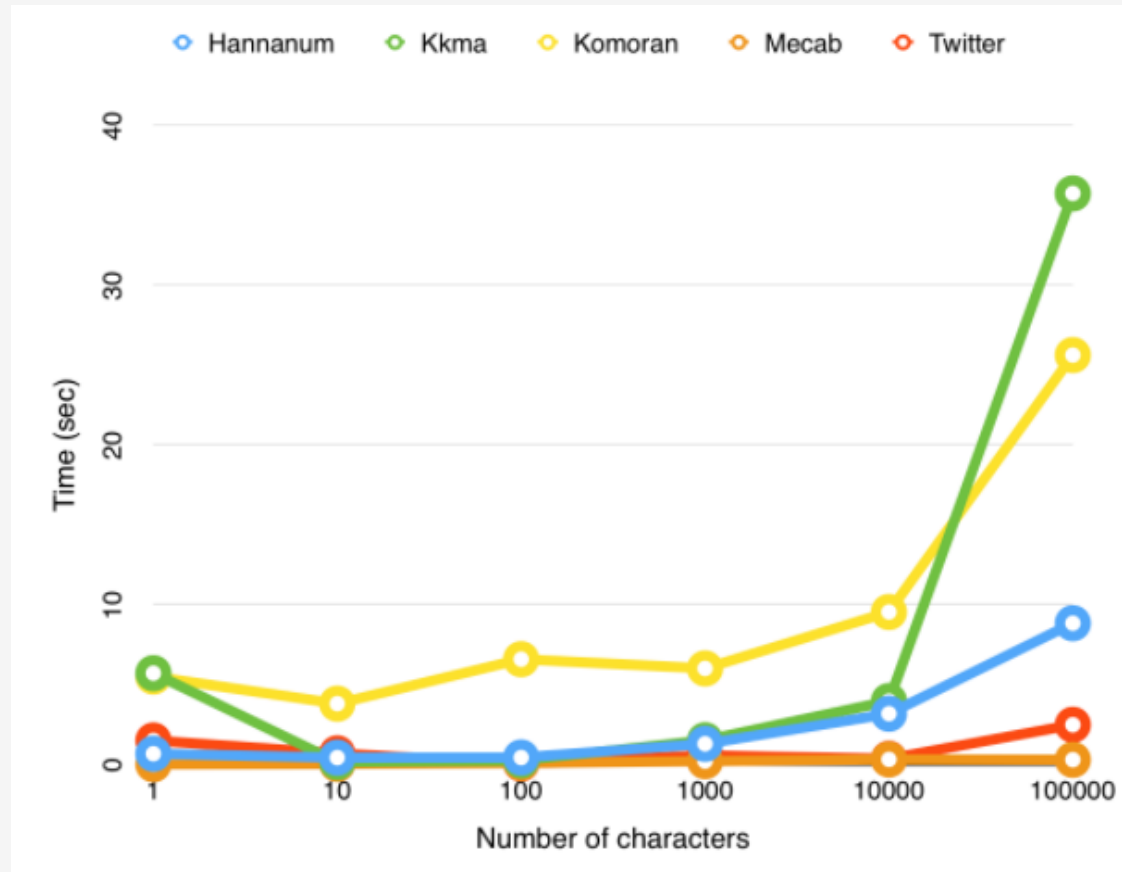
## 형태소 분석 \_ Mecab

```
1 mecab = Mecab()
2 sentence = "제 13회 루빅스 컨퍼런스 발표를 맡은 ToBiz입니다!"
3 temp_X = mecab.pos(sentence)
4 temp_X
```

[('제', 'XPN'),  
( '13', 'SN'),  
( '회', 'NNEC'),  
( '루', 'NR'),  
( '빅스', 'NIP'),  
( '컨퍼런스', 'NNG'),  
( '발표', 'NNG'),  
( '를', 'JKO'),  
( '맡', 'VV'),  
( '은', 'ETM'),  
( 'ToBiz', 'SL'),  
( '입니다', 'VCP+EF'),  
( '!', 'SF')]



## 형태소 분석 \_ Mecab



## 키워드 추출 \_ TF-IDF

개별 문서에서 자주 나타나는 단어에 높은 가중치를 주되, 문서 전반적으로 나타나는 단어에는 페널티를 부과하는 방식의 대표적인 키워드 추출 알고리즘

$$w_{x,y} = tf_{x,y} \times \log \left( \frac{N}{df_x} \right)$$

**TF-IDF**

Term  $x$  within document  $y$

$tf_{x,y}$  = frequency of  $x$  in  $y$

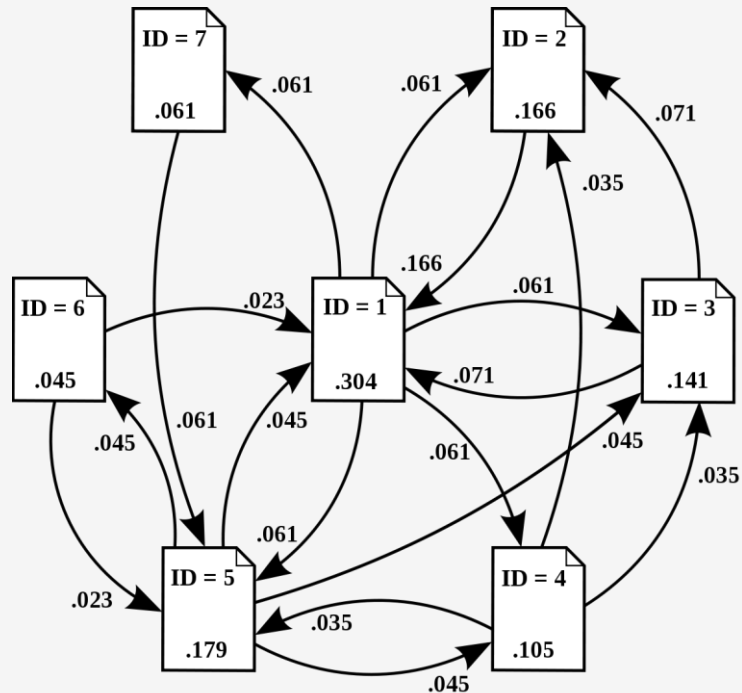
$df_x$  = number of documents containing  $x$

$N$  = total number of documents



## 키워드 추출 \_ TextRank

$$S(V_i) = (1 - d) + d * \sum_{j \in In(V_i)} \frac{1}{|Out(V_j)|} S(V_j)$$



TextRank 는 word graph 나 sentence graph  
를 구축한 뒤, Graph ranking 알고리즘인  
PageRank 를 이용하여 키워드를 추출

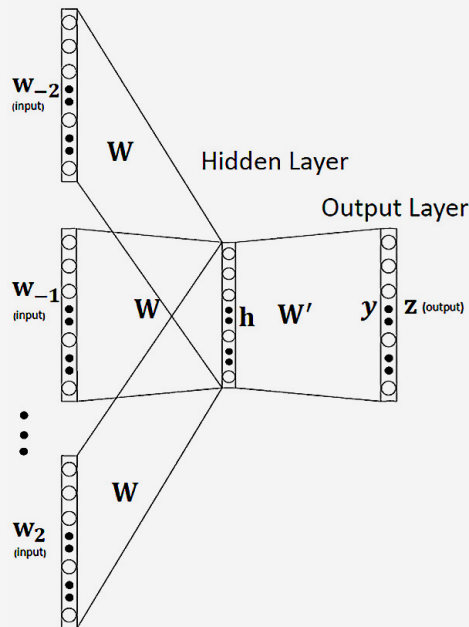
Graph-based ranking 알고리즘은 각 정점의  
정보만을 고려하지 않고 전체 그래프의 글로벌  
정보를 재귀적으로 계산하여 정점의 중요도를  
결정



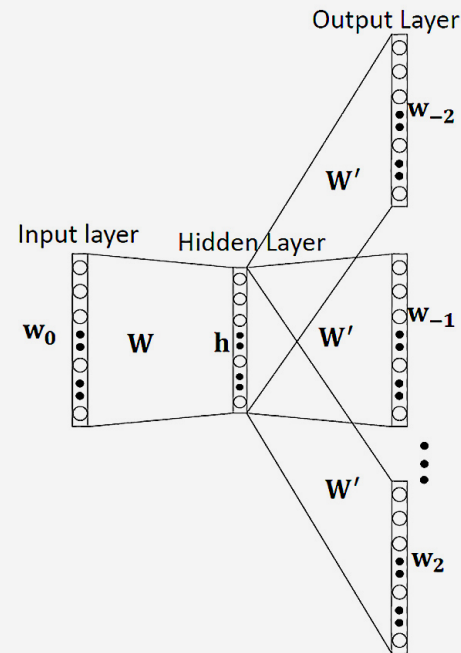
## 자소서 기반 기업 추천 시스템 \_ FastText

**fastText**

Input layer



(a) CBOW



(b) Skip-Gram

FastText의 인공 신경망을 학습한 후에는 데이터 셋의 모든 단어의 각 n-gram에 대해서 워드 임베딩이 가능

데이터 셋만 충분한다면 내부 단어(Subword)를 통해 모르는 단어(Out Of Vocabulary, OOV)에 대해서도 다른 단어와의 유사도를 계산할 수 있는 장점이 존재

잡코리아, 잡플래닛의 모든 합격자소서에 대한  
명사와 동사를 추출하여 학습



최소 5개 이상의 합격 자소서에서 나온 단어들을 바탕으로  
'Skip-Gram'을 이용하여 100차원의 단어 임베딩 벡터 생성



코사인 유사도를 통한  
자소서 기반 기업별 합격 점수 도출 및 기업 추천 서비스



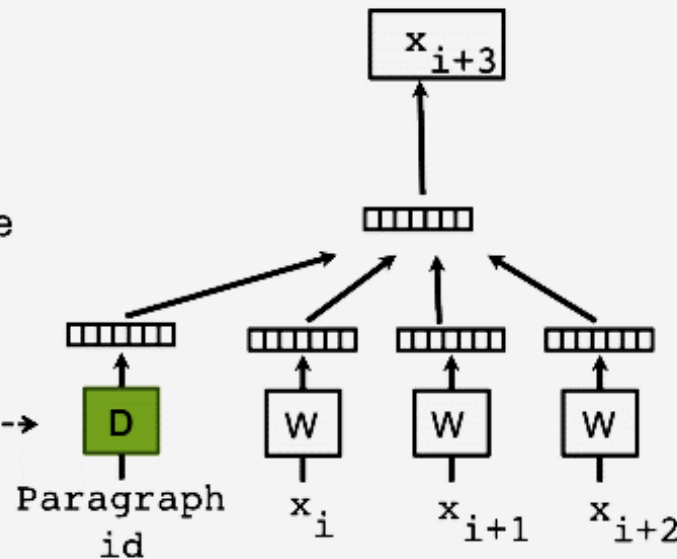
## 자소서 기반 기업 추천 시스템 \_ Cosine similarity

코사인 유사도로 임베딩 벡터 유사도 추출

Cosine similarity

Average/Concatenate

Paragraph Matrix----->



$$\text{similarity} = \cos(\theta) = \frac{A \cdot B}{\|A\| \|B\|} = \frac{\sum_{i=1}^n A_i \times B_i}{\sqrt{\sum_{i=1}^n (A_i)^2} \times \sqrt{\sum_{i=1}^n (B_i)^2}}$$

Word Embedding Vector

Doc2Vec (Vector Average)

Cosine Similarity

Company Recommend (Scoring)



# 유사도를 기반으로 측정된 SCORE 값 기반 기업&직무 추천



To\_biz  
@whatismyjob



Job Opening



Resume



Company

Question.

자기소개를 해주세요. (1500자)

"데이터 분석에 대한 열정과 다양한 경험"

통계, 머신러닝 등 데이터와 관련된 많은 분야를 경험하였습니다. 전공인 통계학과에서 다양한 통계 이론과 데이터를 다루는 방법들에 대해 배울 수 있었습니다. 그러나 통계학 지식만 가지고 산업에 접목시키기엔 많은 부족함이 있다고 생각을 하여 머신러닝 분야에 대해 스스로 공부하였습니다.

또한, 이론만으로 그치지 않고 교내 빅데이터 학회에 들어가 다양한 프로젝트를 진행하였습니다. kaggle의 데이터 셋뿐만 아니라 웹에 존재하는 데이터를 크롤링하기도 하며 다양한 종류의 데이터를 다루며 배운 방법론들을 적용시켜 볼 수 있었습니다. 4학년이 되었을 땐 더욱 현업에 가깝고 수준 높은 프로젝트를 경험하고자 OOO 연구실에 들어가게 되었습니다. 학부 수업과 병행하는 연구실 생활은 많은 어려움이 있었습니다. 프로젝트를 진행을 위해 석박사분들은 기본지식으로 여기던 기법들에 대해서 저는 스스로 공부해야 했습니다. 회의 내용을 놓치지 않으려 회의록을 정리하며, 프로젝트에 조금이라도 더 참여하기 위해 주말을 가리지 않고 밤낮으로 공부하였습니다. 이를 통해 이중 데이터 처리방안, 자연어 처리 등 학부과정에서 다루지 않는 다양한 기술을 배울 수 있었으며, 많은 프로젝트를 경험할 수 있었습니다. 이러한 경험과 열정을 밑거름으로 하여 카카오가 제공하는 다양한 서비스에서 만들어지는 데이터로부터 가치를 창출해내는 데이터 사이언티스트가 되고자 합니다.

..

Showing Top 10 Recommendation Result

자소서를 기반으로 추천된 회사와 직무를 점수와 함께 보여줍니다.

1	LG유플러스	빅데이터·AI(인공..	87.6%	공고
2	롯데쇼핑	빅데이터·AI(인공..	81.2%	공고
3	라인플러스	빅데이터·AI(인공..	76.5%	공고
4	카카오게임즈	빅데이터·AI(인공..	72.3%	공고
5	한화생명보험	빅데이터·AI(인공..	71.4%	공고
6	카카오	빅데이터·AI(인공..	68.2%	공고
7	한국세라믹기..	분석지원(무기계약..	57.7%	공고
8	대한항공	리서치·통계·설문	52.4%	공고
9	롯데멤버스	마케팅·광고·분석	35.0%	공고
10	한국교통연구원	마케팅·광고·분석	32.0%	공고



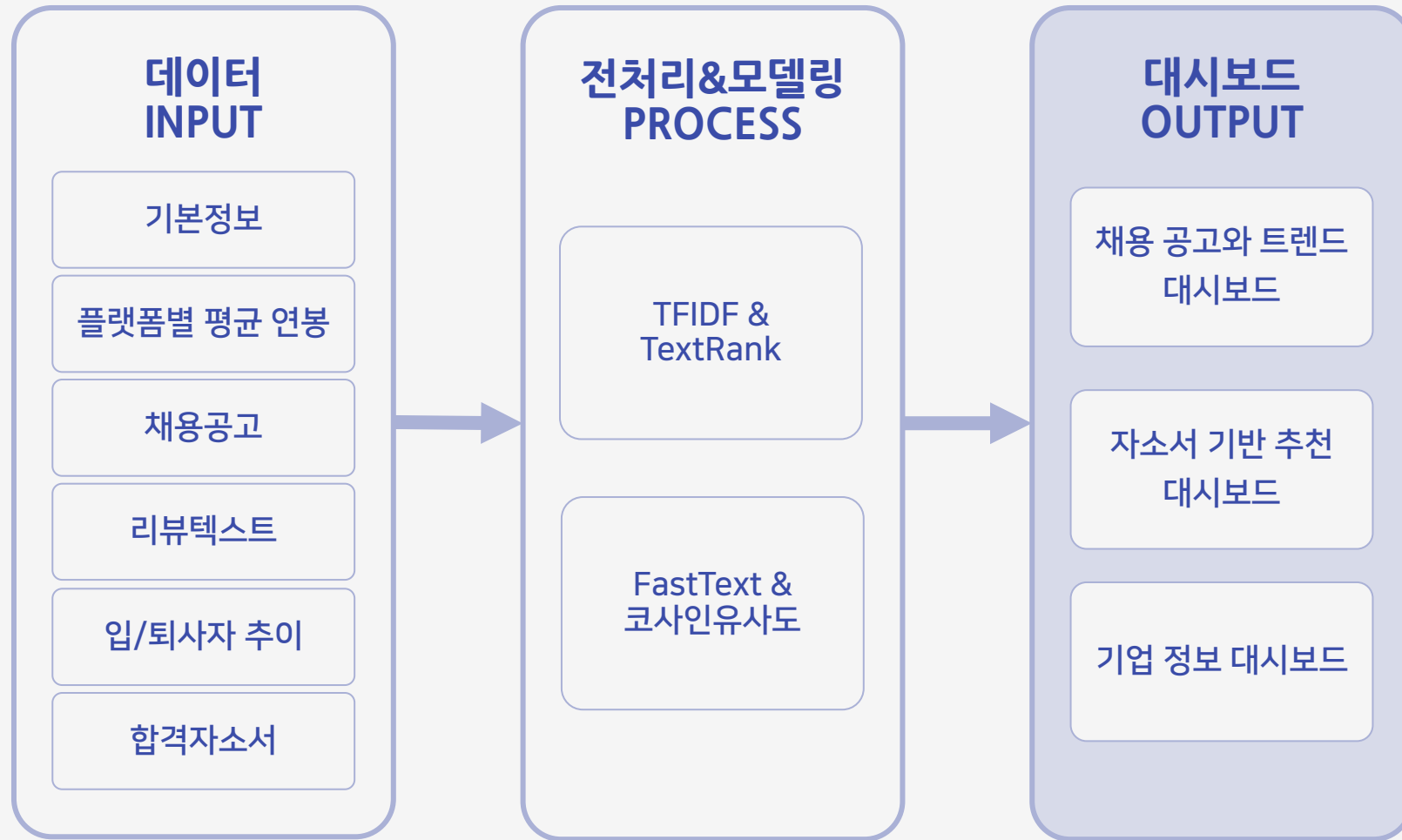
# Dashboard

대시보드 설명 & 시연








## WHERE WE AT?




# 대시보드 설명 & 시연

To\_biz ▼

Job Opening

Resume

Company

LOG-IN

ID

@whatismyjob

Desired Job

PM(프로젝트 매니저)


☐ 로그인 유지


로그인


© 2022 All rights reserved.




## 대시보드 설명 & 시연

To\_biz

Job Opening

Resume

Company

LOG-IN

ID

@whatismyjob

Desired Job

PM(프로젝트 매니저)

☐ 로그인 유지

로그인

© 2022 All rights reserved.



# CONCLUSION

기대효과



## 기대 효과

---

01.

취업 단계별 필요정보  
효과적 습득

02.

급변하는 채용시장  
적응에 도움

03.

효율적인 취준

04.

기업별 정보 빠르게 확인



# TEAM MEMBERS

15기



김동현



이수민



조준혁

16기



김송민



김주호



정수연

감사합니다

