

CONTENT

- 1 팀 구성원 역할 및 과정
- 2 주제 선택 배경 (INTRO)
- 3 데이터 구축
- 4 모델학습 결과
- 5 이미지 전처리
- 6 서비스 흐름도
- 7 기대효과

Part 1 구성원 역할 및 과정



김도연

공윤성

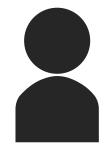
박혜림

송기라

지리산

조성호













- 팀장
- 개발 담당
- 데이터 전처리 및 모델학습
- DB 및 웹서버 구축
- 자료조사
- 총괄 담당

- 개발 담당
- 데이터 전처리
- 데이터구축
- 통계자료 구성
- 자료조사
- 이미지 처리 알고리즘

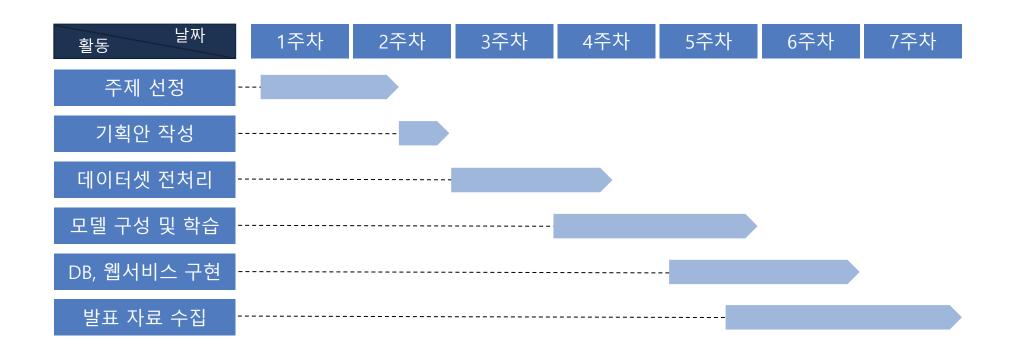
- 개발 담당
- 데이터 전처리
- 데이터구축
- 통계자료 구성
- 자료조사
- 데이터 시각화

- 보조개발 담당
 - 크롤링 담당
- PPT 구성
- PPT 디자인
- 자료조사

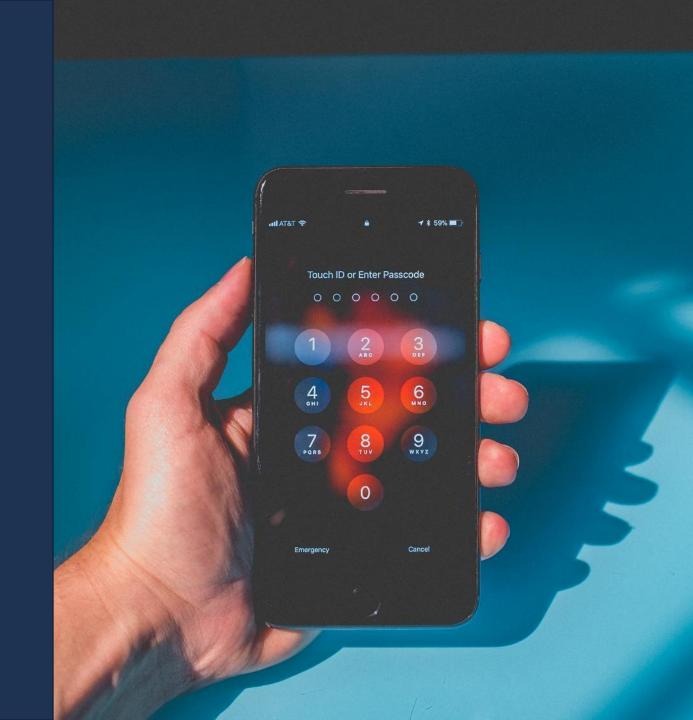
- 보조개발 담당
- 발표 담당
- · 자료조사
- 통계자료 구성

- 보조개발 담당
- 데이터 구축
- 자료조사
- 통계자료 구성

Part 1 과정 소개



Part 2 주제 선택 배경



프로젝트 동기

세금계산서 신고시 문제점

- 1. 홈택스 사용을 어려워함
- 2. 양이 많으면 신고하기 힘듦
- 3. 문서화 작업하는 데 시간이 오래 걸림
- 4. 누락분 발생 확률이 높음



프로젝트 목표

대상

주로 (종이•수기) 세금계산서를 발급•수취하는 사업자

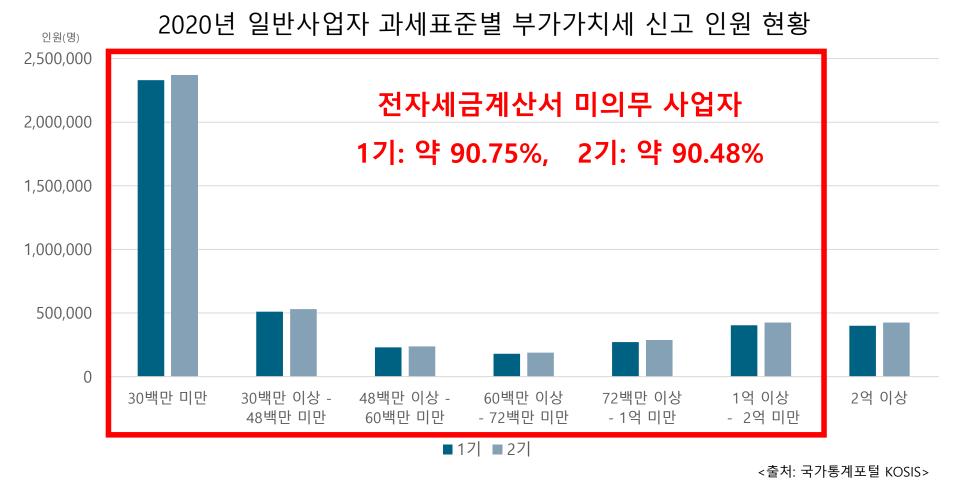
방법

핸드폰을 이용해 손글씨 부분을 인식하여 필요 정보 데이터화

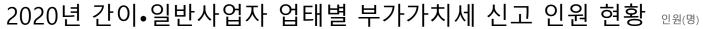
서비스

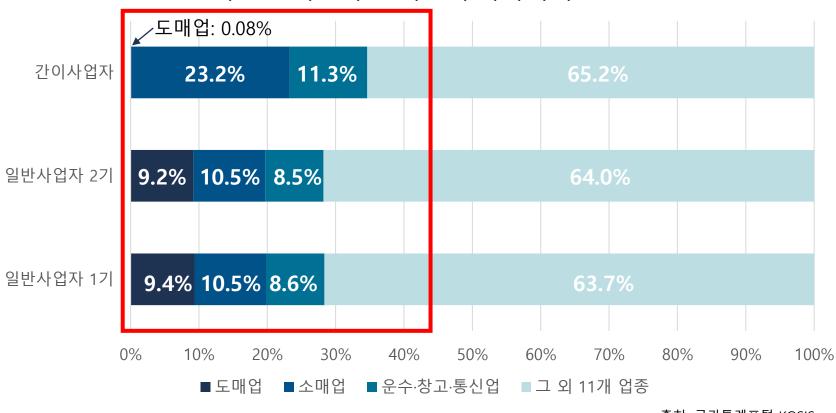
정보들을 데이터 베이스에 저장하고 저장한 데이터 기반으로 세금 신고 간소화

사업자 현황



사업자 현황





동대문 의류 도소매

	섬유패션 전체	동대문 패션시장
매출(원)	87조	15조 (전체의 17%)
수출(조)	14.7조	3.5조 (전체의 21%)
업체(개)	23.4만 제조,도소매 총합	2.5조 (전체의 11%)
종사자(명)	68.6만 제조,도소매 총 고용	18.1조 (전체의 26%)

<출처: 문화체육관광부·동대문관광특구협의회·업계 종합, MarIT>



현재 동대문에서 옷들을 사입하고 있는데요.

세금계산서를 발행해야 할 것 같은데 요청하면 보통 메일로 보내주나요? 아니면 혹시 제가 부가세를 입금을 해주어야 하는 건가요?



동대문은 한 달에 한 번 정도 올라가고 사입은 사입삼촌들이 해주고요. 물건 올 때 <mark>장</mark>끼를 받는데 이건 세금을 안 내게 해주는 건 아니라고 들어서... 세금계산서를 받아야 한다고 들었거든요.

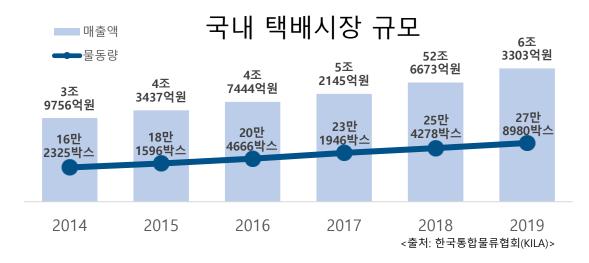
거래명세서(장끼)는 비용 인정이 가능할까?



No!

거래 내용을 나타내는 문서일 뿐 세금 신고시에 비용 처리가 되지 않아 세금 폭탄의 원인이 된다.

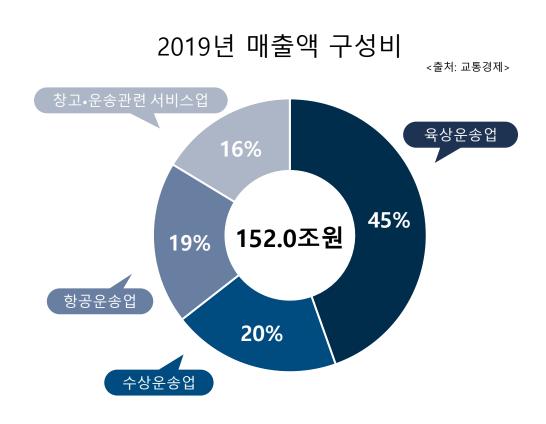
물류 운수업



기업체 수 및 매출액

구분	(개, 십억원)	2014년	2015년	2016년	2017년
전체	기업체 수	178,080	179,075	181,627	184,559
선세	매출액	31,078	31,620	33,560	34,900

<출처: 국가통계포털 KOSIS, 2019.6.7. 기준>



수기 세금계산서를 많이 사용하는데 규모가 커짐에 따라 신고 문제가 빈번해짐

농수산물 도소매

농가 인구 추이



한국농어민신문

유통 > 유통현장

2007.08.02

"현금영수증 발급, 걱정되네"

|농산물 도매·재래시장, | 5000원 안팎 소액결제·60대 이상 고령상인 많아

신고 포상제 악용 '전문 신고꾼 활개 우려

<출처: E-나라지표>

<출처: 한국농어민신문>

영세사업자 세금신고의 어려움

부가가치세 신고 납부 기한 (일반과세자, 간이과세자)

사업자구분	기분	법정신고기한	제출대상서류
	1기예정	4. 1 ~ 4. 25	1. 부가가치세(예정 또는 확정)신고서 2. 매출처벌세금계산서합계표 3. 매입처벌세금계산서합계표 IDH대 함목은 해당하는 경우예만 제출] 4. 역세을 첨부서류
일반과세자	1기확정	7. 1 ~ 7. 25	5. 대손씨액공제신고서 6. 매입세액 불공제분 계산근거 7. 매출처발계산처합계표 8. 매입처발계산처합계표 9. 신용카드매출전표등수행명세서 10. 전자화폐결제명세서(전산작성분 첨부가능)
204414	2기예정	10. 1 - 10. 25	11. 부동산임대공급가백명세서 12. 건물관리명세서 13. 한금매출명세서 14. 주사업장 총괄납부를 하는 경우 사업장별 부가가치세과세표준 및 납부세백(환급세백) 신고 명세서 15. 사업자단위과세를 적용받는 사업자의 경우
	2기확정	1. 1 ~ 1. 25	15. 사업자전취 보세를 작용받는 자업자의 승무에는 사업자단위과세의 사업장별부가가치세과세 표준 및 납부세액(환급세액)신고명세서 16. 건물등감가상각자산취득명세서 17. 의제매입세액공제신고서 18. 그 밖의 필요한 증빙서류
간미과세자	1.1~12.31 (1년)		1. 부가가치세(예정 또는 확정)신고서 2. 매입처벌세금계산서합계표 [마래 항목은 해당하는 경우에만 제출] 3. 매입자발행세금계산서합계표 4. 업체물 첨부서류(업세물 해당자) 5. 부동산임대급자 6. 사업장(한공급) 학명세서(부동산임대업자) 6. 사업장(한공급) 사이 7. 의제매입세액공제신고서 8. 그 밖에 『부가가치세법 시행규칙』 제74조에 따른 해당 서류

〈출처: https://movieranking.co.kr/archives/326〉



라이프 > 생활정보

2020, 12.16

영세사업자를 위한 간이과세, 어떻게 달라지나?

간이과세자의 납부할 세액을 계산하는 것은 조금 다르다. 그냥 매출액에 업종별 부가가치율을 곱하고 그 금액에 10%를 곱하여 매출세액을 계산하고, 공제세액을 뺀다. 공제세액으로는 매입세금계산서 등 수취세액, 의제매입세액, 매입자발행 세금계산서 세액, 전자신고세액, 신용카드매출전표 등 발행세액이 있다.

업종별 부가가치율은 다음과 같다.

전기・가스・증기 및 수도사업: 5%

소매업, 재생용 재료수집 및 판매업, 음식점업: 10% 제조업, 농·임·어업, 숙박업, 운수 및 통신업: 20% 건설업, 부동산임대업, 그 밖의 서비스업: 30%

INTRO

매년 바뀌는 세법









<출처: 대한민국 정책브리핑/기획재정부>





帝日新聞

2021.05.06

세무서 " 대면 신고 안 받아요"... 홈텍스 어려운 자영업자들 '답답'

| 코로나로 대면 창구 운영 중단...회원가입·인증 과정 애 먹어 | 세무사사무실 대행하면 수수료 발생...'이중과세' 주장도 | 제도적 보완 필요...안정된 납세 환경 만들어야

긍부정 분석

소셜 크롤링한 '세금계산서 ' 를 긍부정 분석한 결과

부정 단어 비율이 높음.





블로그-뉴스 분석 21.10.04~21.11.03 <세금계산서+이슈>



LDA 크롤링데이터(지식인, 뉴스기사)

전체 지식iN (1-10/5,702)

✓ 정확도 ✓ 최신순 ✓ 추천순

기간 전체 ▼

면세 영세 사업자 세무사답변 2021.10.10.

영세사업자는 부가세 신고를 하고 면세 사업자는 사업장 현황신고를 해서 면세사업자는 매입세액을 공제 받을 방법이 없는건가요? 부가세 신고를 하지못해서...?? 면세...

Q&A > 부가가치세 | 답변수 1 · 추천수 0

영세/과세신고 뭐가 다른 건가요? 세무사답변 2021.08.13.

... 사업자 입장에서 영세을 신고 > 결론 : 부가세 환급 과세 신고 > 결론 : 부가세 환급 이렇게 이루어지는 것 아닌가요?ㅠㅠ 영세 신고는 대표적으로 수출하는 사업자만...

Q&A > 부가가치세 | 답변수 1 · 추천수 0

영세한 1인 개인사업자에요. 현재 간 2021.09.08.

영세한 1인 개인사업자에요. 현재 간이사업자이고 개인작업실에 사업자내고 종종... 워낙 영세하고 작은곳이라 기장료도 내야하는지, 신고 때만 비용을 내면되는지...

1		제목	본문	답변
2	0	면세 영세	영세사업지	안녕하세요 세무사 유민식 입니다부가세가 과세되는 하는 사업을 하
3	1	영세과세신	안녕하세요	안녕하세요 매출 매입 일 경우영세매출이라면 매출세액 매입세액
4	2	영세한 1인	! 개인사업	간이과세자의 경우는 연매출 만원미만이라면 매입과 상관없이 부가;
5	3	추가내공 '	영세 상품	제가 틀리게 해석하고 있는게 있다면 댓글부탁드립니다영세가 아닌
6	4	긴급고용인	내일 화요	언제부터인가 네이버 광고글 땜시 지식IN인 아니고 광고IN 이 되어기
7	5	아버지 영	아버지가	안녕하세요네이버지식 세무분야의 다크호스 세무사 스머프 라고합니
8	6	영세사업자	사업개업한	사업자이더라도 소득증명이 가능하시면정부지원 대출 자격이 충족이
9	7	근로소득자	안녕하세요	안녕하세요 추가납부 세금은 없습니다 사업을 하고 있는 것이 맞다면
10	8	영세사업자	영세 사업기	안녕하세요 네이버지식인 전문세무상담위원 신선호세무사입니다영세사업;
11	9	간이과세지	간이과세지	간이과세자는 일반과세자 등 년간 매출액이 만원에 미달하는 영세사업자를
12	10	코로나 영	안녕하세요	홈택스에서 출력 가능합니다국세청 홈텍스라고 네이버에 검색해보서

크롤링 주제

사람들이 부가세나 세금 신고 에 관련하여 어떤 생각이 주를 이루는지 다양한 분야에서 정보를 크롤링

- 뉴스: 세금계산서, 수기 세금신고
- 지식인: 세금신고 어려움, 세금계산서 불편함, 영세사업자신고
- 조사와 한글자 단어를 불용어 처리 후 프로그램 시행
- 분량
 - 지식인 : 주제마다 약 2000개의 데이터 수집
 - 뉴스기사 : 주제마다 약 3000개의 데이터 수집
 - 네이버뉴스 세금계산서 기사.csv
 - 네이버뉴스수기 세금 신고저장.csv
 - 🔊 세금신고어려움지식인 크롤링.csv
 - 📭 세금계산서불편지식인 크롤링.csv
 - ☞ 영세사업자신고지식인 크롤링.csv

유사 사례



뉴스 > SW

2021.10.13

화물맨, 편의성 극대화 위해 하반기 대규모 업데이트

화물맨은 기존 업무 처리의 비효율을 개선하기 위해 무료 전자 세금계산서를 발행 기능을 제공할 방침이다.

이전까지 화주들은 차주에게 운송료 지급을 위해 종이 세금계산서를 직접 받아 엑셀로 일일이 수기로 입력했고, 차주들은 운송완료 건에 대한 계산서를 작성, 송부하는 과정을 반복해야 했다.



2021.07.19

디지털 막차 탄 운송물류업계... 종이 세금계산서, 전자 계산서만큼 간편 처리

운송물류업계도 PDA, 스마트폰, REID 등 다양한 정보기술이 적용되고 있지만 영세 개인사업자들은 유독 세금계산서 발행에 종이 계산서를 고집하는 경향이 있다.

운송/주선업체를 대상으로 종이계산서로 인한 비효율을 제거하는데 초점을 맞춰저 있는 것이 특징이다. 반복적인 수작업에 의존하던 화물, 운송업계의 운송료 정산 업무를 AI기술과 데이터로 자동화했다.

Part 3 데이터 구축



데이터 구축

AI-HUB한글문자OCR활용

다양한 형태의 한글 문자 OCR				
소개	다운로드			
데이터셋명	데이터셋명 다양한 형태의 한글 문자 OCR			
데이터 분야	음성/자연어	데이터 유형		이미지
구축기관	동양시스템즈		담당자명	천성열(엔티정보)
가공기관	엔티정보	데이터 관련 문의처	전화번호	043-225-7890
검수기관	엔티정보		이메일	csy78990@gmail.com
구축 데이터량	117만	구축년도		2020년
버전	1.0	최종수정일자		2021.06.27

구분		데이터 수	
인쇄체		75,000	
필기체	글자	344,412	
2/1/1	단어	756,813	
합계		1,176,225	

Al-Hub "다양한 형태의 한글문자 OCR"

- * 알고리즘 성능 개선을 위한 손글씨 등의 다양한 형태의 한글문자 이미지 데이터 셋 구축
- * 필기체 : 성별, 나이 연령대가 다른 152명 작성자가 작성한 한글
- * 인쇄체 : 실생활에서 사용되는 서식 50종에 대한 개인정보 재가공

데이터셋 가공한 이미지

77777 7+ 7+ 7+ 7+ 7+ 7+ 가 가 가 가 가 가 가 가 기ト 기ト 기ト 기ト 기ト 가 가 가 가 TH TH 가 オ 가 가 가 가 가 アト アト アト アト 71 71 71 71 71 フトフトフトフトフトフトフトフト

네이버나눔손글씨

나눔손글씨 딸에게 엄마가

옛날에는 에모장에 할 말을 써서 냉장고에 붙였 던 추억이 있어요. 아제는 다 커버린 딸이 어쩌면 잇었을지 모르는 얼마의 줄싸체. 나눔손글씨 세아체

방교 초등학교 2학년 2반 3번입니다. 엄마 아빠가 회사에서 제 정성스러운 굴씨체로 행 북하게 일했으면 좋겠습니다. 나눔손글씨 무궁화

대한민권이라는 이름만 들어도 가슴이 벅찬 정찰 준비행 입니다. 우궁화 꼭막채점 일판단심 영원히 대한민국 국인이고 싶습니다.

나눔손글씨 암스테르담

사랑하는 이가 한글을 배우고 있어요 앍기 쉽게 자음과 모음을 써줍니다. 언어와 문화를 넘는 글씨를 나누고 싶어요. 나눔손글씨 성실체

말하지 않아도 나를 나라내는 글씨, 글씨에는 내가 담겨있고, 누군가를 향한 진심을 표현할 수 있는 가장 좋은 도구인 것 같아요. 나눔손글씨 외할머니글씨

매일 한글을 적으며, 잊지 않으려 하시는 92세 되할머니 글씨입니다. 한글을 사랑하는 되할머니 의 마음을 공유합니다.

나눔손글씨 할아버지의나눔

할아버지가 써 보는 한글입니다. 요즘 유행하는 줄씨체는 아니겠지만 누군가에게는 꼭 필요한 글씨체였으면 좋겠습니다. 나눔손글씨 아빠글씨

작업이 운전인 아빠는 손이 많이 거차와다. 그 손으로 멋진 글씨를 쓰시는 우리 아빠! 멋진 아빠 인생 응원하고, 사랑합니다. 나눔손글씨 다행체

더 정성스러운 마음을 전함 때 손글씨를 끄쳐 입니다. 자욱마다 마음의 씨앗이 담기면 좋 겠어요. 당신에게 닿아 꽃이 띄도록.

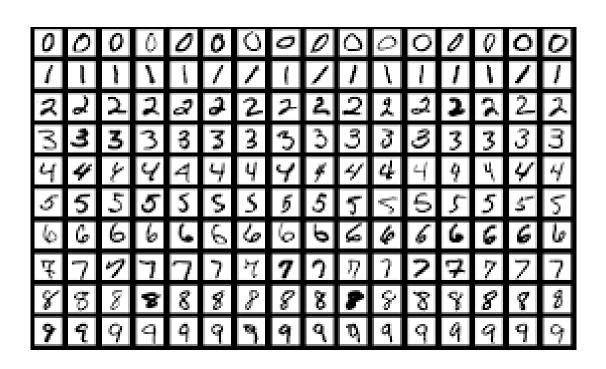
NAVER "나눔손글씨 글꼴"

- OCR 및 이미지 생성 기술을 통해, 256자만으로
 손글씨 고유의 특징을 살려낸 자연스러운 글꼴
- 글꼴 종류 : 109종

<출처: 네이버 클로바>

Part 3 데이터 구축

MNIST-0.17



MNIST-0.17

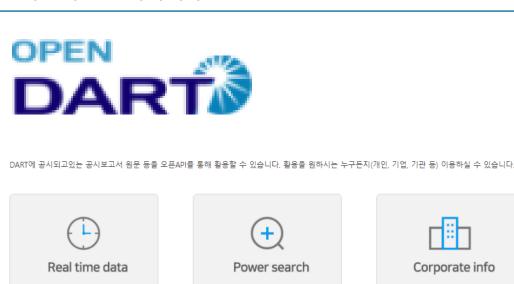
MNIST classifier with average 0.17% error

MNIST-0.17

- 60,000개의 트레이닝 셋과 10,000개의 테스트셋으로 이루어져 있고 28x28 배열과 라벨데이터로 이루어져 있음
- 0~9까지 총 10개의 클래스를 가지고 있으며,
 색상 채널이 없는 흑백 이미지
- 평균 오류 발생률 0.17% 나오는 모델 구성

데이터 구축

사업자등록번호 데이터 수집



DART 공시원문 활용

DART에 공시되는 공시보고서의 원문을 XML형식으로 다운로드받아 원하는 자료를 자유롭게 추출하여 사용할 수 있습니다.

주요공시 및 재무정보 제공

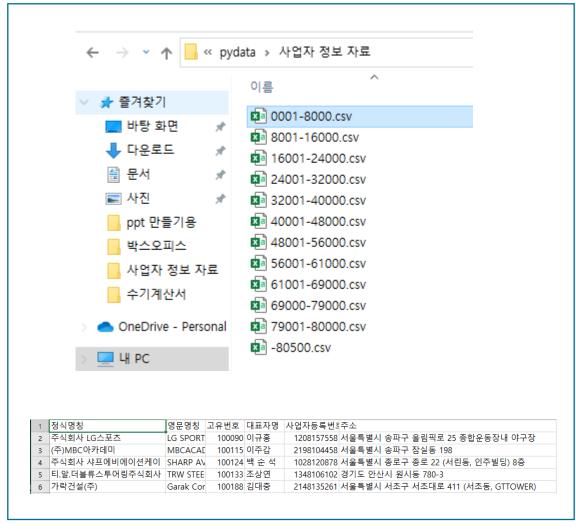
사업보고서 • 증권신고서 • 주요사항 보고서 주요항목 및 주요재무계정, 지분보고서 종합정보를 데이터 형식으로 바로 활용할 수 있습니다.



대용량 재무정보 제공

상장법인(금융업 제외)에서 제출한 전체 재무제표를 분기별로 다운로드 받을 수 있습니다.

- 2020년 1월부터 금융감독원에서 제공하는 전자공시시스템 오픈 API
- 상장기업의 재무제표 정보만 제공



<출처: open Dart> © Saebyeol Yu. Saebyeol's PowerPoint

데이터 구축

직접 수기로 작성한 세금계산서(40여개)

API를 이용하여 수집한 파일의 실제 사업자 등록번호 참고하여 작성

정식명칭	영문명칭	고유번호	대표자명	사업자등록번호	주소
제석주택주식회사		437824	김선용	4138105928	전라남도!
(주)한국대기사	TAIKISHA	437879	츄마세이지	1168154120	서울특별시
리보피아주식회사	Ribopia Co	437912	황성빈	1088155074	서울특별시
태성산업개발(주)	TAESUNG	437958	황규백	1308171862	경기도 안
(주)쓰리디컴넷	3dcomnet	437967	김종길	4028126530	전라북도
(주)이너텍	Innertech	437994	이현호	1388116693	경기도 수
(주)히타치국제전기	HITACHI k	438027	시노모토	34837	AKIHABAF
팅크웨어(주)	Thinkware	438036	이흥복	2298126731	경기도 성
(주)바이넥스홀딩스	Binex Hold	438063	정명호	2148683676	서울특별시
(주)퍼시픽유니온	PACIFIC U	438230	윤영술	1208615005	경기도 용역
유비에스캐피탈비브이	UBS CAPI	438319	Jan derk B	16874	경기도 용역
(주)중앙캐피탈	Joong Ang	438328	이경훈	1148605744	서울특별시
(주)에너지월드	energy wo	438337	추 교 환	3148150747	대전광역시
(주)이건리빙	EAGON LI	438425	최명래	1218155849	인천광역시



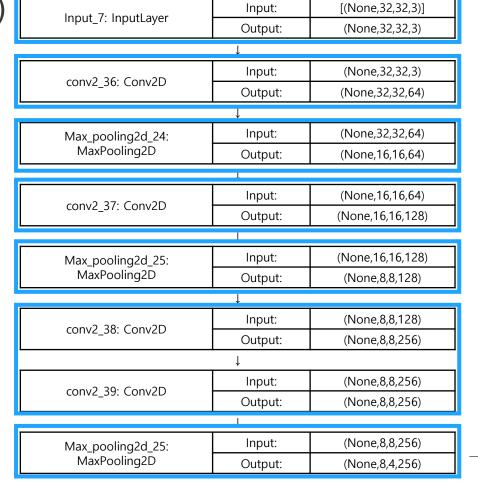
<출처 : OPEN DART>

Part 4 모델학습 결과



모델학습

최종 모델(한글)



500 / 2 40: Con / 2D	Input:	(None,8,4,256)
conv2_40: Conv2D	Output:	(None,8,4,512)
	↓	
batch_normalization_12:	Input:	(None,8,4,512)
BatchNormalization	Output:	(None,8,4,512)
	Ţ	
conv2 41: Conv2D	Input:	(None,8,4,512)
CONV2_41. CONV2D	Output:	(None,8,4,512)
	Ţ	
batch_normalization_13:	Input:	(None,8,4,512)
BatchNormalization	Output:	(None,8,4,512)
	Ţ	
Max_pooling2d_27:	Input:	(None,8,4,512)
MaxPooling2D	Output:	(None,7,2,512)
	<u>I</u> .	
flatten_6: Flatten	Input:	(None,7,2,512)
natten_o. Flatten	Output:	(None,7168)
	ļ	
drangut 6: Drangut	Input:	(None,7168)
dropout_6: Dropout	Output:	(None,7168)
	Ţ	
danca 6: Danca	Input:	(None,7168)
dense_6: Dense	Output:	(None,2350)

Part 4 모델학습

최종 모델(숫자)1

Input_22: InputLayer	Input:	[(None,28,28,1)]			
input_22. inputLayer	Output:	(None,28,28,128)			
↓					
conv2 190: Conv2D	Input:	(None,28,28,128)			
conv2_189: Conv2D	Output:	(None,28,28,128)			
	1				
2 100 Can 2D	Input:	(None,28,28,128)			
conv2_190: Conv2D	Output:	(None,28,28,128)			
	1				
2000 2 101. Com 2D	Input:	(None,28,28,128)			
conv2_191: Conv2D	Output:	(None,28,28,128)			
	↓				
batch_normalization_63:	Input:	(None,28,28,128)			
BatchNormalization	Output:	(None,28,28,128)			
global_average_pooling2d_84:	Input:	(None,28,28,128)			
GlobalaveragePooling	Output:	(None,128)			
	↓				
reshape_63: Reshape	Input:	(None,128)			
resnape_03. Resnape	Output:	(None,1,128)			
	↓				
dense_147: Dense	Input:	(None,1,128)			
delise_147. Delise	Output:	(None,1,4)			
	↓				
dense_148: Dense	Input:	(None,1,4)			
dense_140. Dense	Output:	(None,1,128)			

Input:	(None,1,128)
Output:	(None,28,28,128)
↓	
Input:	(None,28,28,128)
Output:	(None,28,28,128)
↓	
Input:	(None,28,28,128)
Output:	(None,28,28,128)
↓	
Input:	(None,28,28,128)
Output:	(None,28,28,128)
↓	
Input:	(None,28,28,128)
Output:	(None,28,28,128)
Ţ	
Input:	(None,28,28,128)
Output:	(None,128)
+	
Input:	(None,128)
Output:	(None,1,128)
↓	
Input:	(None,1,128)
Output:	(None,1,4)
<u></u>	
Input:	(None,1,4)
Output:	(None,1,128)
	Output: Input: Output: Input: Input: Output: Input: Output:

모델학습

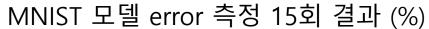
최종 모델(숫자)2 🏻

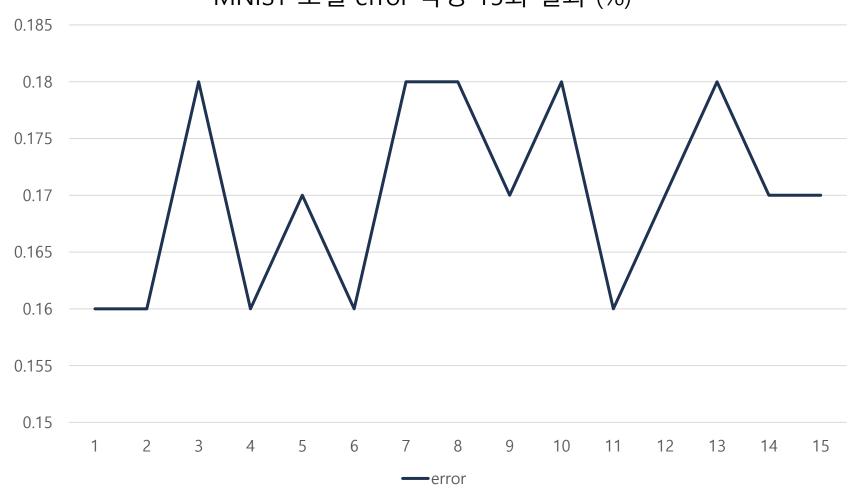
Multiply_64: Multiply	Input:	(None,1,128)		
Multiply_64. Multiply	Output:	(None,28,28,128)		
average_pooling2d_42:	Input:	(None,28,28,128)		
AveragePooling	Output:	(None,14,14,128)		
	<u></u>			
conv2_195: Conv2D	Input:	(None,14,14,128)		
CONV2_193. CONV2D	Output:	(None,14,14,128)		
	<u> </u>			
conv2_196: Conv2D	Input:	(None,14,14,128)		
CONV2_190. CONV2D	Output:	(None,14,14,128)		
	↓			
conv2_197: Conv2D	Input:	(None,14,14,128)		
CONV2_197. CONV2D	Output:	(None,14,14,128)		
	+			
batch_normalization_65:	Input:	(None,14,14,128)		
BatchNormalization	Output:	(None,14,14,128)		
global_average_pooling2d_65:	Input:	(None,14,14,128)		
GlobalAveragePooling	Output:	(None,128)		
	<u> </u>			
rochano 65: Pochano	Input:	(None,128)		
reshape_65: Reshape	Output:	(None,1,128)		

dance 151. Dance	Input:	(None,1,128)
dense_151: Dense	Output:	(None,1,4)
	↓	
danca 152: Danca	Input:	(None,1,4)
dense_152: Dense	Output:	(None,1,128)
	↓	
Multiply_65: Multiply	Input:	(None,1,128)
ividitiply_63. ividitiply	Output:	(None,14, 14,128)
	1	
average_pooling2d_43:	Input:	(None,14, 14,128)
AveragePooling	Output:	(None,7, 7,128)
	↓	
Global_Max_pooling2d_25:	Input:	(None,7, 7,128)
GlovbalMaxPooling2D	Output:	(None,128)
	1	
global_average_pooling2d_65:	Input:	(None,128)
Global Average Pooling	Output:	(None,128)
	1	
concatenate_21:	Input:	(None,128)
Concatenate	Output:	(None,256)
	↓	
dense_153: Dense	Input:	(None,256)
delise_133. Delise	Output:	(None,10)

Part 4 모델학습 결과

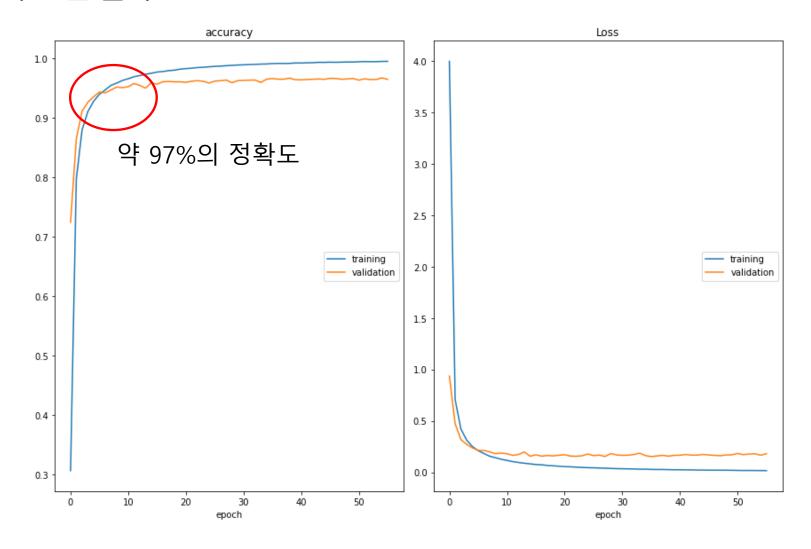
MNIST 숫자 인식 모델 결과





모델학습 결과

한글 손글씨 인식 모델 결과



Part 5 이미지 전처리



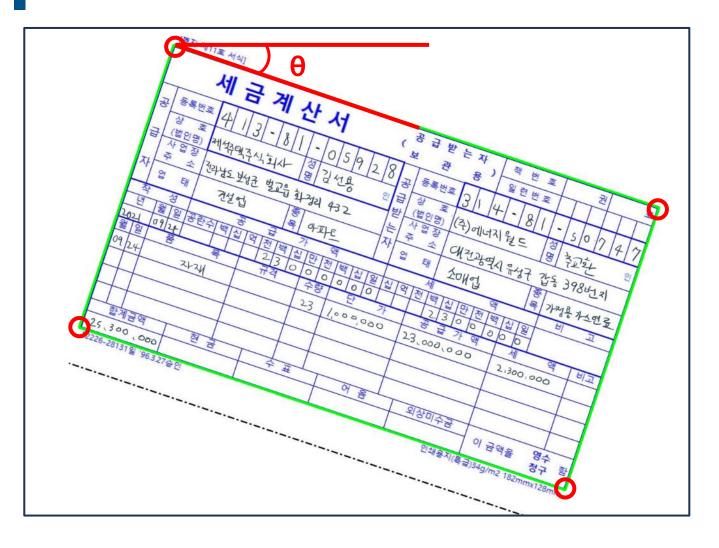
처리할 이미지



※ 조건 ※

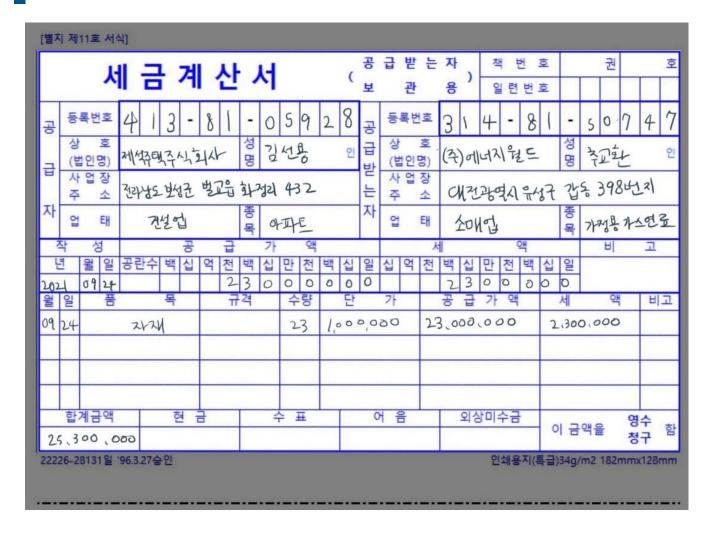
- 1. 제공하는 가이드라인 내 세금계산서가 위치함
 - 2. 양식의 외곽선이 변형되어 있지 않음 Ex) 구겨짐

양식 외곽선 찾기



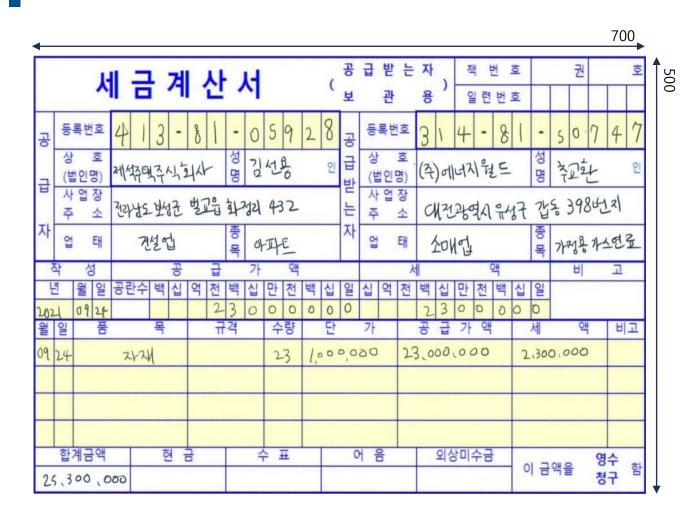
컨투어에 외접하며 면적이 가장 작은 직사각형을 구하고, 기울어진 각도와 네 꼭지점의 좌표값을 얻음

회전 여부 확인 후 필요 영역 추출



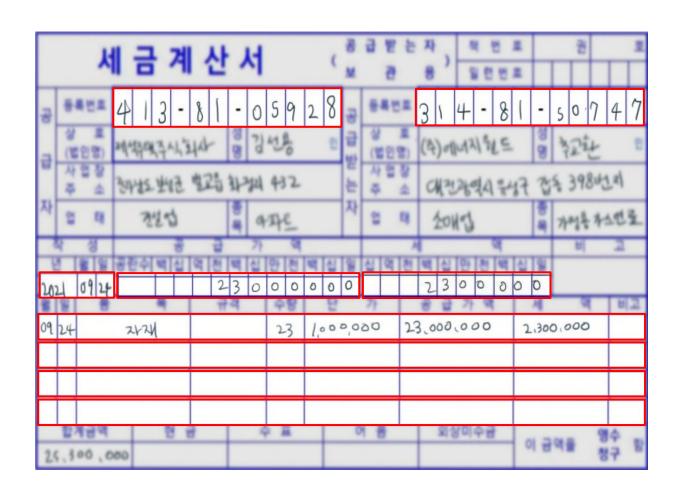
원하는 형태(=회전이 없는 형태)로 변환시키고 필요한 영역만 잘라내기

구간 바운딩박스 처리



700*500 크기로 resize 후 미리 양식에 맞게 데이터가 들어올 칸의 좌표를 미리 설정해 둠

이미지 잘라내기



양식에 맞게 미리 좌표값을 사용해 지정해 둔 영역 부분 추출

이미지 전처리

이미지 잘라내기



한 번 추출했던 부분을 각각 박스별로 분할 추출하고, 공란과 '-'를 제외한 글씨를 순서를 부여하여 숫자 정보가 바뀌지 않도록 설정함

이미지 전처리

모델 예측 후 리스트화 - 숫자

```
확률값
              예측값
[[1.2149426e-04.3.2045698e-06.7.8389830e-06.2.5649630e-05.6.1774324e-07.
  6.1090327e-07 1.1138483e-04 1.6565461e-07 9.9972206e-01 6.9231837e-06]]
값은 : 8
[[1.0448982e-09 1.4789602e-07 2.8434039e-07 9.9999583e-01 8.7865809e-10
  2.1043888e-07 7.7790844e-08 1.2717872e-07 2.6986470e-06 5.9669543e-0711
값은 : 3
[[4.62668424e-07 9.99959826e-01 5.06634592e-07 1.44051855e-08
  2.75123493e-05 1.30446938e-06 7.67628171e-06 2.58919272e-06
  9.51252019e-08 7.51015392e-0811
값은 : 1
[[3.2641984e=07.9.3304102e=08.2.4696816e=09.9.3277059e=08.9.9962723e=01
  1.9547878e-07 1.1621389e-07 7.5801972e-06 3.2063751e-04 4.3647367e-051
값은 : 4
[[5.2796537e=07.3.1035249e=06.9.9996173e=01.9.3481704e=06.1.8631406e=06
  6.4622157e-08 1.2854065e-05 8.1229300e-06 2.2237650e-06 1.2420801e-0711
값은 : 2
```

한 글자씩 자른 이미지를 MINST-0.17 모델을 이용하여 예측하고 리스트로 저장

Part 5 이미지 전처리

모델 예측 결과 - 숫자

손글씨	4138105928	3148150747	2300000	730000
모델 예측 결과	4138105928	3148150747	23600000	2300000
정답율	10 / 10	10 / 10	7 / 8	6 / 6

이미지 전처리

모델 예측 후 리스트화 - 한글

```
예측값
확률값
filename: '0.jpg', <mark>confidence: 0.9996,</mark> string: '건설업'
filename: '1.jpg', <mark>confidence: 0.1356,</mark> string: '울픔'
filename: '2.PNG', confidence: 0.6340,
string: '세종특별자치시 도움5로 19 정부세증청사 8동'
filename: '3.jpg', <mark>confidence: 0.9975,</mark> string: '임태'
filename: '4.jpg', <mark>confidence: 0.3956,</mark> string: '왐규 백'
```

단어 혹은 문장의 글씨 이미지를 예측하고 리스트로 저장

Part 5 이미지 전처리

모델 예측 결과 - 숫자

손글씨	건설업	罗玉	세종특별자치시 도움5로 정부세종청사 8동	임대	智和叫
모델 예측 결과	건설업	울픔	세종특별자치시 도움5로 19 정부세증청사 8동	임대	왕규 백
정답율	3 / 3	1 / 2	18 / 19	2/2	2/3

Part 6 서비스 흐름도



로그인/메인

사업자 등록번호를 ID로 활용하여 로그인



서비스 흐름도

자료입력 / 내 자료 조회

메인화면에서 각각 자료입력 / 내 자료 조회를 선택시 화면



현재 로그인 사업자번호: 3148150747

공급가액: 23000000 세액: 2300000 날짜: 2021-11-03

상대면세여부 : 과세사업자

매입/출 여부 : 매입

상대사업자번호: 4138105928

공급가액: 23000000 세액: 2300000 날짜: 2021-11-03

상대면세여부 : 과세사업자 매입/출 여부 : 매입

상대사업자번호: 4138105928

공급가액: 23000000 세액: 2300000 날짜: 2021-11-03

상대면세여부 : 과세사업자

매입/출 여부 : 매입

상대사업자번호: 4138105928

메인으로 돌아가기

서비스 흐름도

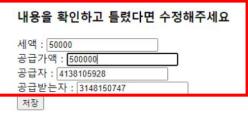
OCR결과 / 업로드 후 자료조회

이미지 업로드 후 OCR결과확인 / 저장후 자료 확인

			人	I	3	겨	1	산		서				(공보	급	받관	는	자용)		_	번	io io			권			Đ.	
공	등	록번	ė	4	3	-	8	1	-	0	5	9	2	8	공	등	록반	立	3	1	4	-	8	1	-	5	0	7	4	7	
급		인		제4	택주	٨	44		성명	김	也	e o		인	급받		인		(4) ot	47	119	اع		성명	なて	Pź	t		인	
	사주	집	장 소	હ્યા	5.坐	记	艺	74	弘	겧리	4:	32	9		는	사주	업	장 소	0	N전	광	택시	1 to	17	갑	50	398	841	21		
자	61		EH	-	전설이	d		중 아파트							자	6.0		태	1	201	12				寄号	가	77 B	1	소연로		
	낙	ć		1	V a	공	7	급		가		액		Va de				Å				1	액	The second			비		ī		
202		8	일	공란4	는 백	십	억	전 2	3	십	만	천	0	십	0	십	억	천	백	십	만	천 o	0	십	일						
월	일		普	2	목			T	24		수	량		단		가			공	급	가	Q			세		Qi	1	비	ı	
09	24			24	M						2	-3	13	0 0	٥٥	00		23	3.0	00	0	00		2	.30	0.0	000)			
20		계급 00		වට	Š	ē ;	3			-	-	#			(H f	el .			외성		수금	}	C	1 =	액	2	명정		햠	

.....

OCR 결과



취소

현재 로그인 사업자번호: 3148150747

공급가액: 23000000 세액: 2300000 날짜: 2021-11-03 상대면세여부: 과세사업자 매입/출 여부: 매입

상대사업자번호 : 4138105928

공급가액: 23000000 세액: 2300000 날짜: 2021-11-03 상대면세여부: 과세사양

상대면세여부 : 과세사업자 매입/출 여부 : 매입

상대사업자번호: 4138105928

공급가액: 23000000 세액: 2300000 날짜: 2021-11-03 상대면세여부: 과세사업자 매입/출 여부: 매입

상대사업자번호: 4138105928

공급가액 : 500000 세액 : 50000 날짜 : 2021-11-03 상대면세여부 : 과세사업자 매입/출 여부 : 매입

상대사업자번호 : 4138105928

메인으로 돌아가기

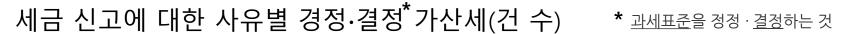
S

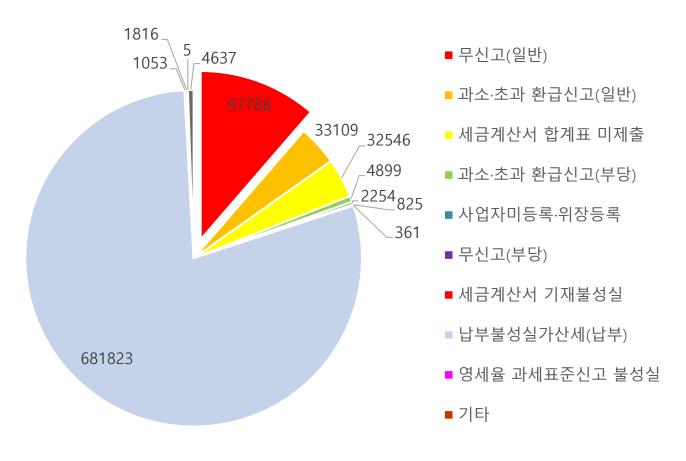


Part 7 기대효과



가산세 방지





<출처: 국가통계포털 KOSIS>

Part 7 기대효과

SWOT분석

- 일반인들의 세무처리 비용 절감
- '삼쩜삼' 비즈니스 모델 성공 사례
- 잠재 고객 폭이 넓음

S

- '삼쩜삼' 을 제치고 시장 장악 가능성
- 정부지원사업의 가능성

- 직접적으로 금전적인 부분이기에 accuracy가 높아야 함
- 각기 다른 영수증을 식별 능력

W

• 세무,회계사들의 반발 가능성

• 대기업 인수 가능성

Part 7 기대효과

SWOT분석

- 늘어나는 시장에서 필요성이 증가해가는 분야인 만큼 높은 정확도를 살려 정부지원 사업 대상이 될 수 있다는 점을 강조
- 세법 개정으로 수기 계산서의 사용자가 증가 할 때 간단하게 사용할 수 있는 모델로서 접근

■ 높은 정확도를 바탕으로 세무,회계사들에게 도움이 될 수 있다는 점을 강조

다른 분야의 플랫폼 보다 사용이 쉽다는 점을 강조

SO

WO

ST

- 양식이 다른 영수증에 대한 대응을 추가해 정부지원 사업 중에 새로운 양식의 영수증이 생겨도 대응 할 수 있다는 점을 강조
- 다른 영수증을 추가하면서 새로운 수기 계산서의 사용자가 새롭게 사용하는 영수증에 대응

- 약간의 오차로 큰 금액의 차이가 생길 수 있는 만큼 세무, 회계사들이 도울 수 있는 분야가 있다는 점을 강조
- 실제 사용사례를 추가해가며, 오차가 생기는 일을 줄여 다른 플랫폼 보다 안정적인 모델로

Part 7 기대효과

사업확장



- (1) S 유형 성실신고대상자 (복식부기의무자)
- (2) A 유형 외부조정대상자(복식부기의무자)
- (3) B 유형 자기조정대상자(복식부기의무자)
- (4) C 유형 복식부기의무자
- (5) D 유형 간편장부대상자
- (6) E 유형 타소득이 발생한 단순경비율 적용 대상자(간편장부대상자)
- (7) F 유형 단순경비율 대상자(납부세액 발생자, 간편장부대상자)
- (8) G 유형 단순경비율 대상자(납부세액이 발생하지 않은자, 간편장부대상자)
- (9) I 유형 성실신고 사전 안내자 (복식부기 또는 간편장부)
- (10) V 유형 주택임대소득 분리 과세 대상자

	종합소득세 계산 절차	비고
종합소득금액	총수입금액 – 필요경비 = 소득금액 (사업소득, 근로소득, 이자소득, 배당소득, 연금소득, 기타소득)	
과세표준	종합소득금액 – 소득공제 = 과세표준 (기본공제, 추가공제, 연금보험료공제, 특별소득공제, 노란우산공제 등)	
산출세액	과세표준 X 종합소득세율 = 산출세액 (6% ~ 42%)	
세액공제·감면	특별세액공제(보험료, 의료비, 교육비, 기부금 등), 기 장세액공제, 배당세액공제 등	
가산세	무신고, 과소신고, 납부불성실, 증빙불비, 무기장 등	
총결정세액	산출세액 - 세액공제·감면 + 가산세 = 총결정세액	
납부(환급)할 세액	총결정세액 - 기납부세액 = 납부할 세액 (중간예납, 수시부과, 원천징수 등)	

<출처:아이보스>



감사합니다