<도입부>

안녕하십니까.

벤처(Venture)캐피탈에게 올바른 투자 요인을 조언(Advice)해주는 팀.

“어드벤처(AdVenture)”입니다.

혁신적인 모험을 통해 유니콘 기업이 되고자 하는 벤처기업.

그러한 벤처기업에 높은 위험을 감수하며 투자하는 벤처캐피탈에

잘 어울리는 이름이라고 생각하여

팀명을 “어드벤쳐”라고 정하게 되었습니다.

[어드벤쳐]

[AdVenture]

<팀원 소개>

간단한 팀원 소개를 하겠습니다.

팀장은 저 박태원이고, 팀원은 경소연님, 변종우님으로 구성되어있습니다.

<문제 정의>

최근 코로나바이러스와

전쟁과 같은 전세계적인 이슈들로 인해

자산가치와 물가는 롤러코스터를 탄 듯 대호황과 대공황을 오가는

극단적인 변동성을 보여주었습니다.

[사진1]

2020년 코로나 팬데믹으로 인해 디지털 전환은 가속화되었지만

실물 경제는 급격히 침체하였습니다.

이에 각국 정부는 얼어붙은 소비 심리를 활성화하기 위해

유동성을 공급하였으나,

이 자금들이 의도와는 다르게 대부분 자산시장으로 유입되어

주식과 부동산 가격의 폭등을 초래하였습니다.

[사진2]

[사진3]

또한 글로벌 생산 기지인 중국의 락다운으로 인해

재화 전반의 가격이 오르고,

유럽에서 벌어진 러시아 - 우크라이나 전쟁,

물류대란 등 여러 요인들이 물가 상승을 가속화시켰습니다.

이러한 '슈퍼인플레이션' 현상에 대응하기 위해

미국 연방공개시장위원회는 사실상 제로금리에 가까웠던 기준금리를

최근 5.25%까지 몇번의 빅스텝과 자이언트 스텝으로 급격히 인상하였습니다.

사진[4]

사진[5]

사진[6]

사진[7]

이렇게 금리가 상승하면서 재무적인 관점에서

고위험-고수익의 성격을 가진 벤처 투자는 큰 위기를 맞게 되었습니다.

금리 상승으로 인해 예금/적금이나 채권 등 안정적인 투자 수익이 보장되면서 투자자들은 더이상 높은 위험을 감수할 필요가 없게 된 것입니다.

이처럼 경기가 어려울 때일수록 펀더멘탈을 중시하는 투자가 주목받게 되는데

이것을 벤처 투자에도 동일하게 적용해 보면 어떨까 하는 생각에

이번 프로젝트의 주제를 ‘벤처 기업 투자 요인 분석 및 예측’으로

선정하게 되었습니다.

주식 투자에 관심이 있다면 ‘기업의 펀더멘탈을 보고 가치투자 하라’ 라는 말을 들어보신 적이 있을 것입니다.

[펀더멘탈][투자]

[Fundamental]

: ‘본질적인’, ‘필수적인’, ‘기본 원칙’

: 기업의 성장 가능성, 영업이익, 매출, 재무상태 등으로 판단할 수 있는 가치를 분석하여 투자하는 방법

상장기업의 경우 매출, 비용, 자산, 부채와 같은 재무건전성과

환율, 금리, 고용률과 같은 거시 경제지표라는 가시적인 펀더멘탈이 있지만

벤처기업은 대체로 그렇지 못합니다.

벤처기업에게는 어떤 다른 요소가 펀더멘탈이라 부를 수 있는 판단지표인지

저희 어드벤처는 그 답을 ‘**중소벤처기업부의 벤처기업정밀실태조사보고서’**에 수집된 데이터에서 찾아보고자 합니다.

<타임라인>

저희의 타임라인입니다.

현재 데이터 수집 과정을 마친 후

데이터 전처리 과정을 진행중입니다.

저희의 타임라인을 보시면 전처리 기간이

굉장히 길게 설정되어 있는 것을 보실 수 있습니다.

먼저 저희 데이터는 2016년에 2049개의 표본과 593개의 컬럼,

2017년에 2114개의 표본과 594개의 컬럼,

2018년에 2059개의 표본과 594개의 컬럼,

2019년에 2052개의 표본과 352개의 컬럼,

2020년에 2500개의 표본과 357개의 컬럼

2021년에 2500개의 표본과 385개의 컬럼으로 이루어져 있었습니다.

처음 데이터를 선정했을 당시에는 19-21년도의 데이터만 가지고 분석을

진행하려고 했으나 표본 데이터의 수가 너무 적었고,

타겟 변수로 설정한 벤처캐피탈 투자 경험 여부 또한

총 6225개의 표본 데이터 중 622개로 약 9:1의 비율을 갖고 있어

16-18년도의 데이터도 추가하여 분석을 진행하기로 하였습니다.

<데이터 수집>

데이터 수집처는 공공데이터포털로 중소벤처기업부의 벤터기업정밀실태조사보고서 데이터를 연도별로 수집하였습니다.

데이터 형태는 설문조사 형식의 데이터입니다.

현재 전처리 후 12385개의 표본과 68개의 컬럼으로 이루어져있습니다.

컬럼들의 예시로는 창업자의 최종학력, 전공, 이전 창업 경험 등이 포함된

창업자 정보 관련 정보들과 해당 기업 종사자들의 최종학력,

정규직, 비정규직 여부 등이 포함된 인력 구조 관련 컬럼

해당기업의 지분구조, 매출구조, 재무정보에 대한 컬럼,

보유 산업재산권에 대한 컬럼 등이 있습니다.

현재 이 6개의 데이터를 하나로 합치는 과정을 끝냈으며,

그래서 앞서 말한 12385개의 표본과 68개의 컬럼으로 데이터가 이루어져있다.

아직 결측치나 이상치 제거는 이루어지지 않았으며, 기타 등등

<컬럼 삭제>

데이터전처리 과정에서 우리가 가장 많은 시간을 들인 부분은

컬럼을 비교하고 삭제하는 과정이었다.

어떤 컬럼들을 삭제했는지..

왜 삭제했는지…(결측치가 너무 많은 컬럼도 삭제)

재무 컬럼과 같이 수치형 데이터는 대부분 비율로 변환하여 파생변수 생성 후

원본 컬럼 삭제

기타 등등

컬럼을 선택할 때 선행 논문 참고.

(교수님 추천 논문 & 우리가 선별한 논문)

남은 컬럼들을 어떤 기준으로 남겼으며

어떤식으로 활용할지

창업자(entrepreneur)의 특성

* 창업가가 벤처기업 관련 산업경험, 경영경험 및 창업경험을 보유하고 있다면 VCs가 신생 벤처기업을 훨씬 더 우수하게 평가한다는 것(창업경험을 보유하고 있는 창업자 는 해당 분야에 대한 지식의 활용이 가능하고, 기업가적 프로세스(entrepreneurial process)에 대한 이해를 경험을 통해 보유 하고 있음)
* 창업자가 소유한 능력 측면에서는 기술 및 경영, 전문성 측면에서는 혼합적인 교육 배경을 가진 창업자를 선호

최종학력 컬럼

* 창업자의 높은 학력은 벤처기업의 재무적 성과에 유의한 영향을 주며 창업자의 높은 학벌을 의미하는 명문대가 가지고 있는 엘리트 집단의 네트워크는 그 자체로 높은 사회적자본으로 간주될 수 있다 .

특허권 등 관련 컬럼

* 단순히 개념적 수준의 아이디어 단계보다는 프로토타입의 개발 또는 특허의 취득 단계까지 개발이 이루어진 경우에는 VC의 투자 자금 유치에 긍정적으로 작용

앞으로 어떤 방향으로 분석을 진행할지

모델을 어떤 것을 쓸지

SHAP에 대한 설명

그래프, 라인, 스크린샷, 텍스트이(가) 표시된 사진

자동 생성된 설명

**[사진1]**

텍스트, 스크린샷, 그래프, 도표이(가) 표시된 사진

자동 생성된 설명

[사진2] – (출처)부동산플래닛

라인, 그래프, 도표, 폰트이(가) 표시된 사진

자동 생성된 설명[사진3] – (출처) KOSIS 국가통계포털

텍스트, 스크린샷, 야외, 사람이(가) 표시된 사진

자동 생성된 설명

[사진4] – (출처) http://www.ktnews.com/news/articleView.html?idxno=123453

텍스트, 세계, 스크린샷, 지도이(가) 표시된 사진

자동 생성된 설명

[사진5] – (출처) https://n.news.naver.com/mnews/article/001/0012948270

텍스트, 스크린샷, 폰트, 번호이(가) 표시된 사진

자동 생성된 설명

[사진6]- (출처)http://www.domin.co.kr/news/articleView.html?idxno=1405496

텍스트, 스크린샷, 폰트, 도표이(가) 표시된 사진

자동 생성된 설명

[사진7] – (출처) https://n.news.naver.com/mnews/article/023/0003761667