

핀테크와 스마트 금융

8장 / 11주차 1차시

8.0

학습목표

- ◆ 분야별 빅데이터 활용 사례를 열거할 수 있다.
- ◆ 기업신용등급의 개념과 신용등급예측 모형의 구축 과정을 이해한다.
- ◆ P2P 대출거래의 개념과 채무불이행 예측 모형의 구축 과정을 이해한다.
- ◆ 빅데이터와 인공지능 융합 비즈니스를 설명할 수 있다.

8.1

국내외 금융빅데이터 활용 사례: 고객빅데이터 분석 및 활용

◆ CRM(Customer Relationship Management, 고객관계관리)

- 조직 전체에 산재되어 있는 고객 데이터를 포착하고 통합하는 동시에 이들 데이터들을 분석하여 고객 전략을 수립하기 위한 정보시스템
- 빅데이터 분석이 가장 많이 활용되는 분야로 고객빅데이터 분석을 통해 차별화된 고객서비스를 제공할 수 있음
- 현 고객에 대한 서비스 이용 성향 및 취향 분석을 통해 고객을 세분화하고, 각 고객군에 맞는 맞춤형 전략을 추진

◆ 최근 고객빅데이터 분석

- 향상된 맞춤형 서비스를 위해 추천 시스템(recommendation systems)을 개발하는데 적용됨

◆ 일본 최초의 인터넷전문은행인 재팬넷은행(Japan Net Bank)

- 고객빅데이터를 기반으로 분석을 수행하고 있음

8.1

국내외 금융빅데이터 활용 사례: 고객빅데이터 분석 및 활용

◆ 일본 인터넷전문은행

- 은행이 보유한 개인 및 법인 예금, 투자신탁, 주식 등의 고객 데이터를 수집하고, IT기업의 빅데이터 분석 기술을 활용하여 대출의사결정을 지원함
(주택을 구입할 예정인 고객이 계약 보증금을 마련하기 위해 예금을 늘리는 경향을 포착하여 적절한 시기에 은행이 주택대출 관련 서비스를 추천 및 제안)

◆ 미국 웰스파고

- 고객의 ATM 조작 이력을 토대로 화면에 표시되는 버튼을 고객별로 최적화

◆ 미국의 체이스뱅크(Chase Bank)

- 고객빅데이터를 활용하여 계좌해약 조짐의 패턴을 보이는 고객을 예측하여 특별 관리를 실시함

8.1

금융빅데이터 활용 사례: 마케팅 빅데이터 분석 및 활용

◆ 빅데이터 마케팅

- 빅데이터로부터 마케팅 전략 인사이트를 도출하는 것으로
빅데이터 기반의 소비행동 패턴 분석과 소비 트렌드 분석, 고객 불만 및 경쟁 이슈 분석을 통해
실시간 마케팅 의사결정 지원함

◆ 비자(VISA Inc.) 신용카드사

- 고객의 결제위치 및 시점, 구매품목을 실시간으로 파악하고 목표 고객(target customer)의 구매 이력 및 성향을 고려하여 타겟 마케팅(target marketing)을 실시함
(주유소에서 결제를 마치면 인근 카페의 상품할인쿠폰을 발송)

◆ 스탠다드차타드은행(Standard Chartered Bank)

- SNS를 이용한 타겟 마케팅을 실시간으로 활용하여 고객유치율 및 수익성을 향상시키고 있음

8.1

금융빅데이터 활용 사례: 마케팅 빅데이터 분석 및 활용

◆ 빅데이터 마케팅

- 빅데이터로부터 마케팅 전략 인사이트를 도출하는 것으로
빅데이터 기반의 소비행동 패턴 분석과 소비 트렌드 분석, 고객 불만 및 경쟁 이슈 분석을 통해
실시간 마케팅 의사결정 지원함

◆ 비자(VISA Inc.) 신용카드사

- 고객의 결제위치 및 시점, 구매품목을 실시간으로 파악하고 목표 고객(target customer)의 구매 이력 및 성향을 고려하여 타겟 마케팅(target marketing)을 실시함
(주유소에서 결제를 마치면 인근 카페의 상품할인쿠폰을 발송)

◆ 스탠다드차타드은행(Standard Chartered Bank)

- SNS를 이용한 타겟 마케팅을 실시간으로 활용하여 고객유치율 및 수익성을 향상시키고 있음

8.1

금융빅데이터 활용 사례: 금융투자 빅데이터 분석 및 활용

◆ 블록체인(blockchain)

- 분산형 데이터베이스와 유사한 형태로 모든 구성원이 네트워크를 통해 데이터를 검증 및 저장하여 특정인의 임의적인 조작이 어렵게 설계된 저장 플랫폼

◆ 블록(block)

- 거래 건별 정보가 기록되는 단위
- 블록이 시간순서에 따라 체인(chain) 형태로 연결된 데이터베이스가 블록체인

◆ 블록체인

- 블록의 정보와 거래 내용(거래정보)을 기록하고, 이를 네트워크 참여자들에게 분산 및 공유하는 분산원장

◆ 중국 알리바바 그룹의 인터넷전문은행인 마이뱅크(MyBank)

- 블록체인 기반 기부 플랫폼을 운영하고 있음
- 블록체인 기술을 이용하여 기부금이 어떻게, 어디에, 얼마나 사용되는지 관계자들이 투명하게 확인할 수 있음

8.1

금융빅데이터 활용 사례: 금융투자 빅데이터 분석 및 활용

- ◆ 금융업에 블록체인 플랫폼을 활용하면 금융거래의 운영절차가 간소화되고, 거래의 인증 및 검증 과정에서 중개기관의 역할이 축소되어 청산 및 결제에 필요한 시간이 단축됨
- ◆ 최초 거래부터 모든 거래내역이 기록·공유되기 때문에 거래상대방에 대한 부정 거래와 위험(counterparty risk, 거래상대방위험) 발생이 감소하며, 거래 과정이 실시간으로 통제되므로 규제 및 감독의 효율성을 높일 수 있음
- ◆ 블록체인이 보안성, 투명성, 탈(脫)제3자, 확장성이라는 강점을 가지고 있으므로 금융 분야에서 블록체인의 활용 가능성은 높아짐

8.2

기업신용등급예측 모형 구축 사례: 신용과 신용등급

◆ 신용(credit)

- 장래의 어느 시점에서 그 대가를 지급할 것을 약속하고 현재의 경제적 가치를 획득할 수 있는 능력을 말함

◆ 신용평가(credit valuation)

- 금융기관은 신용대출 대상의 건전성, 자금용도의 타당성 및 상환능력을 정확하게 파악한 후 여신의 가부판단과 신용대출 가부를 결정해야 하는 것을 말함

◆ 신용평가의 측정치인 신용등급(credit rating)

- 채무자의 채무불이행(default) 가능성에 대해 기업의 경영능력, 운영능력, 시장경쟁력 등 다양한 요인이 반영된 미래 현금흐름과 재무적 융통성을 감안하여 일정한 기호로 나타냄

◆ 신용평가기관

- 알파벳 기호로 표현된 신용등급을 원리금 상환능력의 우열에 따라 "+", "0", "-"의 부호, 즉 노치(Notch)를 붙여 세분화하여 발표하고 있음

8.2

기업신용등급예측 모형 구축 사례: 신용과 신용등급

(1) 신용등급평가의 기능과 목적

◆ 신용평가(신용등급평가)의 기능

- 기업의 신용위험에 대한 정보를 투자자에게 제공함으로써
신용정보의 부족 또는 정보의 비대칭성으로 발생할 수 있는 손실로부터
투자자를 보호하고 발행업체의 자금조달을 용이하게 하는 기능을 수행함

◆ 신용등급평가의 목적

- 채권의 원금과 이자가 약속대로 상환될 수 있는 정도로 측정(심사)하여
투자자에게 전달함으로써 이들 정보를 알지 못하여 발생할 수 있는 손실로부터
투자자를 보호하는데 있음

◆ 기업신용평가

- 회사채 신용등급예측(Corporate Credit Rating Prediction)
- 기업부도예측(Corporate Bankruptcy Prediction) 으로 구분됨

8.2

기업신용등급예측 모형 구축 사례: 신용등급의 의미

- ◆ 신용등급(credit rating)의 기본형은 만기가 1년 이상인 장기채권의 신용등급
- ◆ 채권의 신용위험 정도에 따라 Moody's사는 Aaa부터 C까지, S&P사는 AAA부터 D까지의 등급을 부여함
- ◆ **Aaa**나 **AAA**는 채무불이행 위험이 거의 없어 신용도가 최고인 채권에 부여되며, Moody's사의 Caa ~ C등급, S&P사의 C ~ D등급은 채무불이행 사태가 발생한 채권에 부여됨
- ◆ **투자적격등급(investment grade)**: Aaa ~ Baa등급(Moody's사), AAA ~ BBB등급(S&P사)
- ◆ **투기적 등급(speculative grade)**: Ba ~ C등급(Moody's사) BB ~ D등급(S&P사)

8.2 기업신용등급예측 모형 구축 사례: 신용등급의 의미

◆ [표 8-1] 장기신용등급의 기호와 의미

투자적격등급		의미	투기적등급		의미
Moody's	S&P		Moody's	S&P	
Aaa	AAA	최고위 신용도	Ba1	BB+	채무불이행 가능성 있 나 불확실성 존재함
Aa1	AA+	고위 신용도	Ba2	BB	
Aa2	AA		Ba3	BB-	
Aa3	AA-		B1	B+	신용위험 높음
A1	A+	강한 지급능력	B2	B	
A2	A		B3	B-	
A3	A-		Caa1	CCC+	채무불이행 가능성 높음
Baa1	BBB+	적절한 지급능력	Caa2	CCC	
Baa2	BBB		Caa3	CCC-	
Baa3	BBB-		Ca	CC	채무불이행 중 채권 회수 에 중대한 결함 존재
			C	C	
				D	

8.2 기업신용평가와 개인신용평가

◆ 신용평가

- 기업신용평가(corporate credit ratings)와 개인신용평가(personal credit ratings)로 구분됨

▶ 개인신용평가

- 소비자 금융거래에 있어서 개인의 신용도를 종합적으로 평가하는 것으로
거래자의 상환의사와 **상환능력**을 판단할 수 있는 신용정보에 대한 조사 분석 및 평가를 의미함

▶ 기업신용평가

- 기업의 **부도가능성** 및 **파산가능성**을 평가하여 기업 신용위험의 상대적인 수준을 서열화하고,
위험 수준이 유사한 기업들을 동일한 등급으로 계량화한 지표 값을 예측함

◆ 국내외 신용평가기관들은 신용등급의 예측정확성을 높이기 위해

1960년대부터 현재까지 신용평가모형과 신용등급예측모형 관련 연구를 수행하고 있음

8.2 회사채 신용등급평가와 신용등급예측모형

- ◆ 회사채(corporate bond)는 기업이 자금을 조달하는 방법 중 하나로 액면금액, 만기, 금리 등이 명시되어 있는 회사가 발행한 채권임
- ◆ 회사채는 주로 기업의 자금조달 목적으로 발행되지만 시장의 거래를 통해 기업의 채무상환능력 및 신용위험이 회사채 수익률로 반영됨
- ◆ 기업신용등급평가 또는 신용등급평가예측은 주로 회사채 신용등급(corporate bond rating) 평가를 말함
- ◆ 회사채 신용등급평가는 채권의 원금과 이자가 약정대로 지급될 가능성이 어느 정도인가, 즉 채무불이행(default)의 가능성을 평가하여 이를 간략한 문자와 기호로 나타냄
- ◆ 신용평가기관은 채권을 발행한 기업의 경영실적이나 재무상태가 변화함에 따라서 채권의 신용등급을 조정하여 발표하고 있음

8.2 회사채 신용등급평가와 신용등급예측모형

- ◆ 신용등급평가 또는 신용등급예측 연구는 1970년대부터 수행되었고, 주로 재무변수들을 이용하여 회사채 신용등급예측을 시도함
- ◆ 초창기 연구에서는 신용등급예측모형으로 최소자승법에 의한 회귀모형(Ordinary Least Squares, OLS)를 주로 사용함
- ◆ OLS 이후에는 예측모형으로 다변량판별분석(Multiple Discriminant Analysis, MDA)을 주로 사용함
- ◆ 초창기 연구에서는 신용등급예측을 위한 독립변수로 후순위채 여부에 관한 더미변수, 기업규모, 이익의 변동성(순이익 증가율), 부채비율, 이자보상비율, 총자산순이익률(ROI), 자기자본순이익률(ROE) 등이 주로 사용됨

8.2 재무비율분석을 이용한 신용등급예측

- ◆ 신용등급예측은 부도가능성을 평가하여 신용위험의 상대적인 수준을 서열화한 뒤, 위험수준이 유사한 기업들을 동일한 등급으로 계량화한 지표 값을 예측하는 것
- ◆ 금융빅데이터분석에서는 재무비율분석을 통해 기업의 신용등급평가, 부도확률예측, 감사보고서 평가 및 예측, 상장폐지예측, 배당금예측 등을 수행할 수 있음
- ◆ 금융빅데이터분석을 이용한 기업신용등급예측에 필요한 재무비율 변수의 예시를 아래와 같이 제시하고자 함(7가지 항목비율)
- ◆ **유동성비율**에는 유동비율과 당좌비율이 있으며 이들 비율은 높은 상관성을 가지므로 동시에 사용하는 것은 바람직하지 않음

8.2 재무비율분석을 이용한 신용등급예측

- ◆ **레버리지비율(안정성)**에는 부채비율, 이자보상비율, 차입금/자기자본, 차입금/총자본, 차입금/매출액 비율 등이 있으며 차입금 관련 변수(3가지)는 하나를 선택하여 연구변수로 지정하는 것이 바람직함
- ◆ **성장성비율**은 총자산증가율, 매출액증가율, 순이익증가율 등이 있고,
활동성비율에는 재고자산회전율, 매출채권회전율, 총자산회전율 등이 있음
- ◆ **수익성비율**에는 총자산순이익률(ROI), 매출액순이익률, 자기자본순이익률(ROE) 등이 있으며
시장가치비율에는 주당순이익(EPS), 주가수익비율(PER) 등이 있음
- ◆ 마지막으로 **규모변수**에는 자산총계와 자본총계(자기자본) 등이 있으며 이들 변수는 자연로그로 치환하는 것이 바람직함

8.2

재무비율분석을 이용한 신용등급예측

구분	재무비율	산식
유동성비율	유동비율	$= (\text{유동자산} / \text{유동부채}) \times 100$
	당좌비율	$= [(\text{유동자산} - \text{재고자산}) / \text{유동부채}] \times 100$
레버리지비율 (안정성)	부채비율	$= (\text{타인자본} / \text{자기자본}) \times 100$
	이자보상비율	$= (\text{영업이익} / \text{이자비용}) \times 100$
	차입금/자기자본	$= (\text{차입금} / \text{자기자본})$
	차입금/총자본	$= (\text{차입금} / \text{총자본})$
	차입금/매출액 비율	$= (\text{차입금} / \text{매출액}) \times 100$
성장성비율	총자산증가율	$= [(\text{당기 말 총자산} - \text{전기 말 총자산}) / \text{전기 말 총자산}] \times 100$
	매출액증가율	$= (\text{당기 매출액} / \text{전기 매출액}) \times 100$
	순이익증가율	$= [(\text{금기 말 당기순이익} - \text{전기 말 당기순이익}) / \text{전기 말 당기순이익}] \times 100$

[표 8-2] 신용등급예측을 위한 재무비율과 기업규모변수 예시(1)

8.2 재무비율분석을 이용한 신용등급예측

구분	재무비율	산식
활동성비율	재고자산회전율	$= (\text{매출액} / \text{재고자산})$
	매출채권회전율	$= (\text{매출액} / \text{외상매출금})$
	총자산회전율	$= (\text{매출액} / \text{총자산})$
수익성비율	총자산순이익률(ROI)	$= (\text{세후순이익} / \text{총자산(총자본)}) \times 100$
	매출액순이익률	$= (\text{순이익} / \text{매출액}) \times 100$
	자기자본순이익률(ROE)	$= (\text{순이익} / \text{자기자본}) \times 100$
시장가치비율	주당순이익(EPS)	$= (\text{순이익} / \text{발행주식수})$
	주가수익비율(PER)	$= (\text{주가} / \text{주당순이익})$
규모 변수	자산 총계	$= (\text{부채} + \text{자본})$
	자본 총계(자기자본)	$= (\text{자산 총계} - \text{부채 총계})$ $= (\text{납입자본금} + \text{잉여금})$

[표 8-2] 신용등급예측을 위한 재무비율과 기업규모변수 예시(2)

8.2

빅데이터 분석도구를 활용한 신용등급예측모형 구축

(1) 래피드마이너(RapidMiner) 개요 및 설치방법

◆ 래피드마이너(RapidMiner)

- 2001년 독일의 도르트문트 기술대학(Technical University of Dortmund)의 인공지능 단과대학에서 개발한 오픈소스 분석 패키지 소프트웨어
- 데이터 로드 및 변환, 데이터 전처리 및 시각화, 예측 분석 및 통계 모델링, 평가를 비롯한 데이터마이닝 및 기계학습 절차를 제공함
- 래피드마이너는 기본 데이터 분석 업무를 위해 500개 이상의 오퍼레이터(Operators)를 제공함

- ◆ 래피드마이너 홈페이지(<https://rapidminer.com/get-started/>)를 통해 래피드마이너 스튜디오(RapidMiner Studio) 버전 무료 이용

8.2

빅데이터 분석도구를 활용한 신용등급예측모형 구축

(1) 래피드마이너(RapidMiner) 개요 및 설치방법

 [My Account](#) [Profile](#) [Downloads](#) [Licenses](#)

Personal Information (all questions are required)

First name

First name can not be empty

Last name

Last name can not be empty

Phone Number

 ▼

Your phone number is required

Which usage describes you best?

☒ Student
☐ Professor / Educator

Academic email

Organization

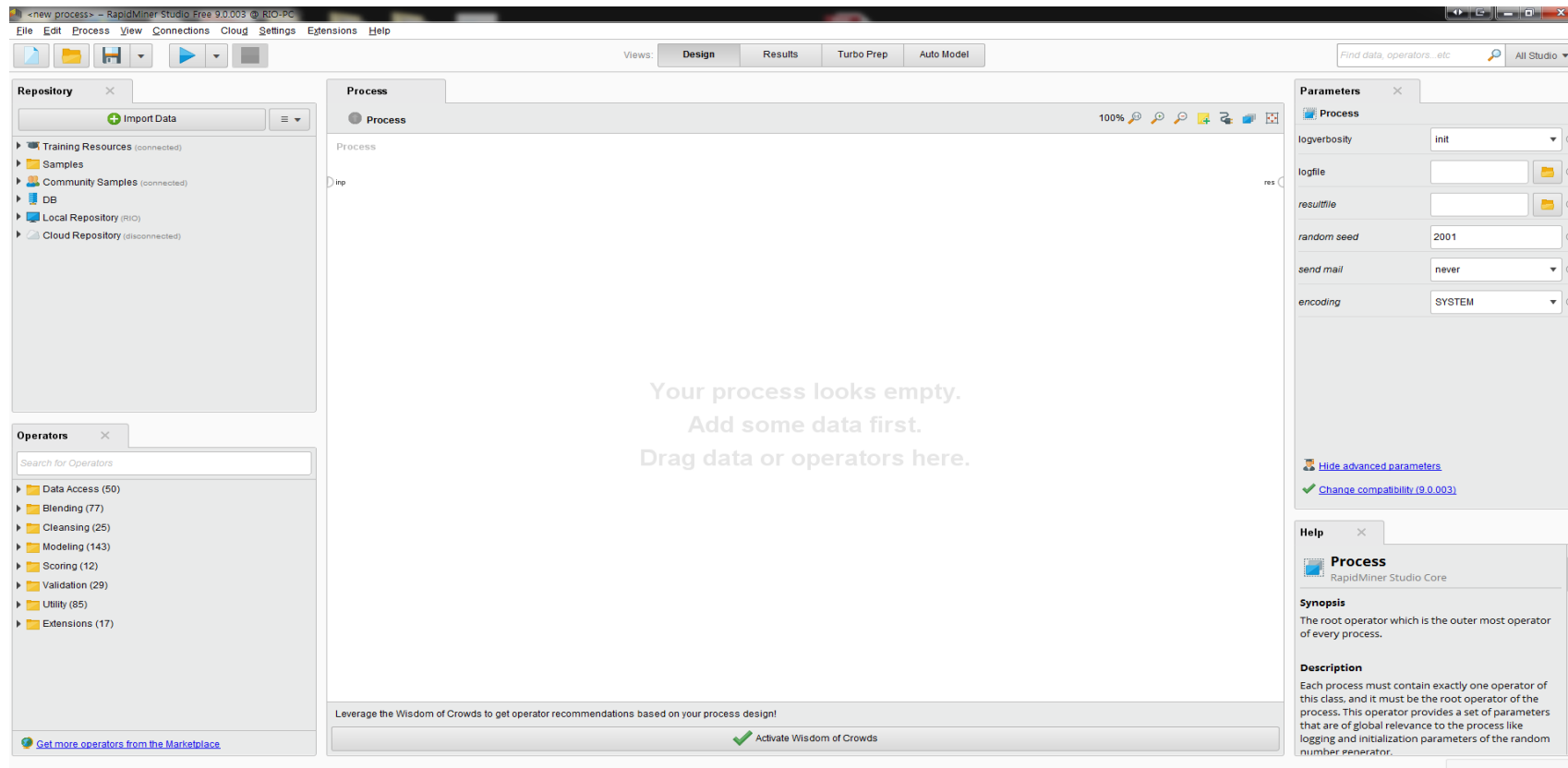
Briefly describe what you will be using RapidMiner for

[그림 8-1] 래피드마이너 교육용 라이선스 신청서 양식

8.2 래피드마이너 사용자 인터페이스 구조

◆ 래피드마이너의 기본 화면

- 저장소(Repository), 오퍼레이터(Operators), 프로세스(Process), 파라미터(Parameters), 도움말(Help) 등의 패널로 구성됨



[그림 8-2] 래피드마이너 사용자 인터페이스 구조

8.2 래피드마이너의 모델링 단계 (1/2)

◆ 래피드마이너의 모델링 단계는 4개의 계층(layer)으로 구성됨

(1) 데이터 읽기

- 오퍼레이터(Operators) 창에 Data Access > Files > Read 폴더 아래에는 CSV, Excel, SPSS, Stata, XML 등 다양한 파일의 읽기를 래피드마이너에서 지원함

(2) 모델 학습

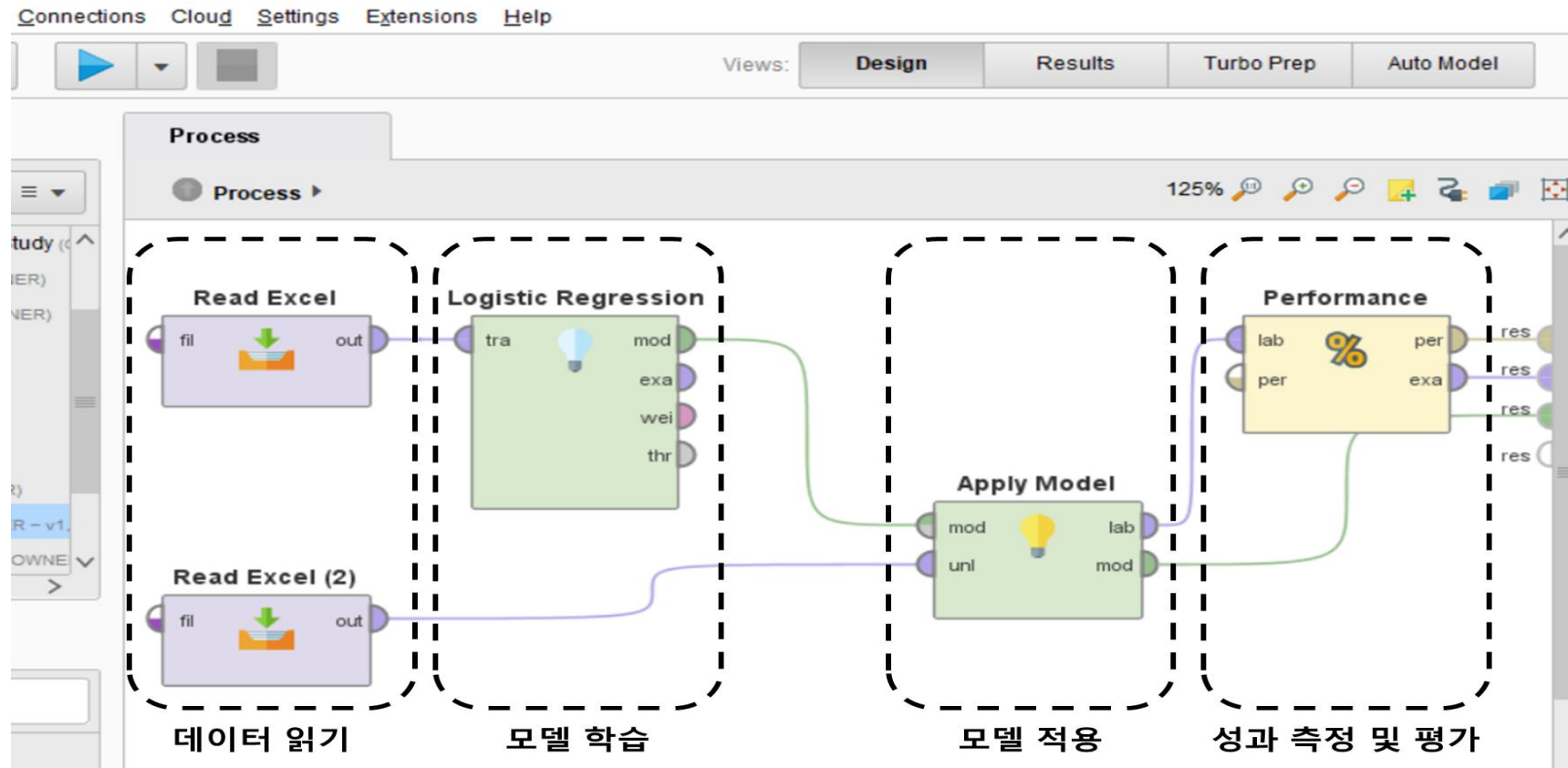
- 데이터 분석 프로세스 설계에서 가장 핵심이 되는 단계로 주요 알고리즘을 선택함
오퍼레이터(Operators) 창에 Modeling 계층 안에는 Predictive, Segmentation, Associations, Correlations 등의 요소가 있음

(3) 모델 적용

- 모델 적용 오퍼레이터(Apply Model)는 모델과 테스트(검증용) 데이터를 입력받아
검증용 데이터 세트의 각 사례에 대해 레이블(종속변수)을 예측하는 역할을 수행함

8.2 래피드마이너의 모델링 단계 (2/2)

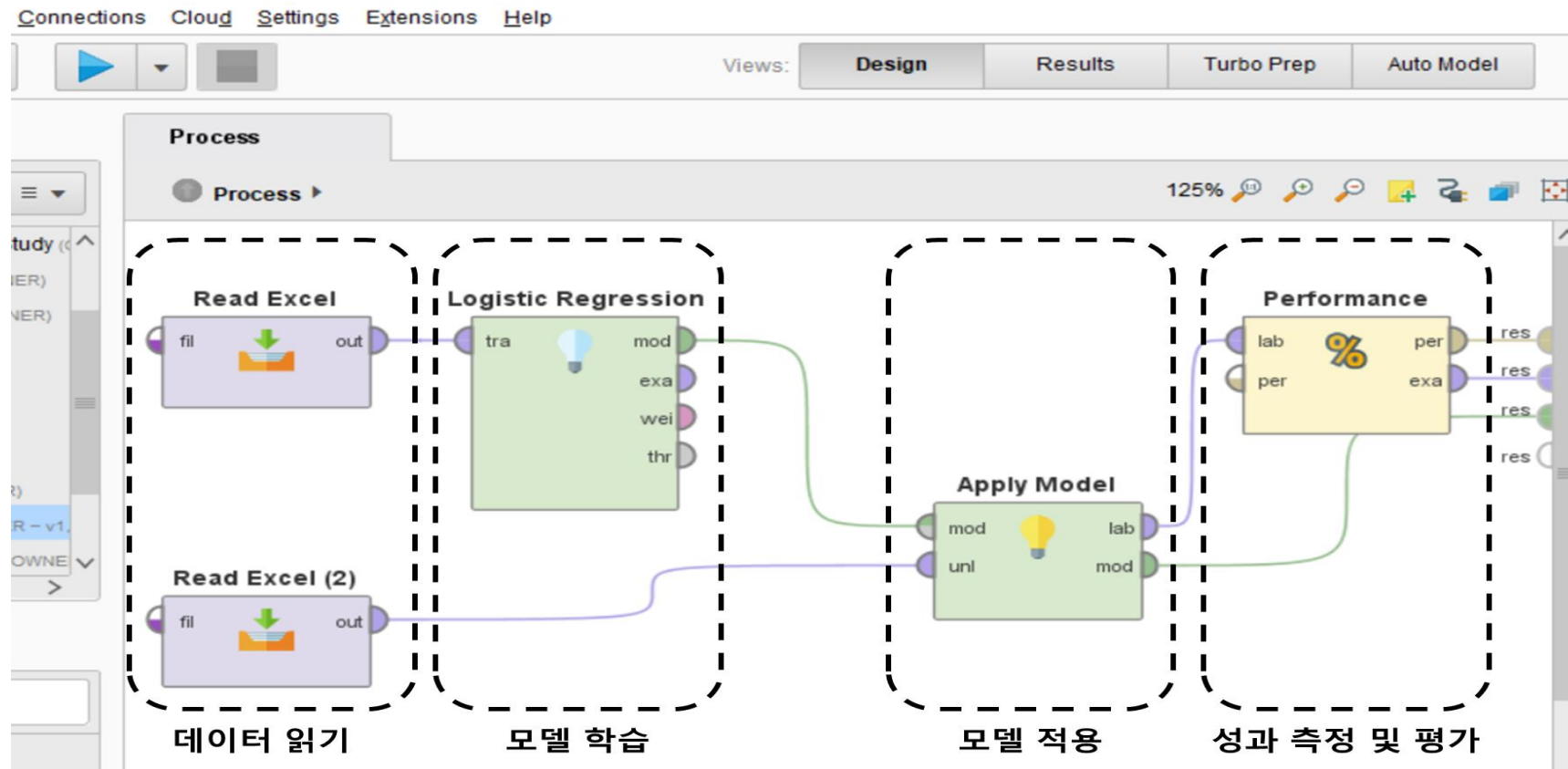
(4) 성과 측정 및 평가: 학습 데이터를 활용해 생성된 모델이 얼마나 잘 작동하는지를 판단하는 것, 성과 평가 방법 중 대표적인 검증 지표는 분류 예측임 - Performance(Classification)



[그림 8-3] 래피드마이너의 모델링 단계

8.2 기업신용등급예측모형 구축 사례

(4) 성과 측정 및 평가: 학습 데이터를 활용해 생성된 모델이 얼마나 잘 작동하는지를 판단하는 것, 성과 평가 방법 중 대표적인 검증 지표는 분류 예측임 - Performance(Classification)



[그림 8-3] 래피드마이너의 모델링 단계

8.2 기업신용등급예측모형 구축 사례

(1) 신용등급데이터 및 연구변수

- ◆ 신용등급예측모형 구축을 위한 실습데이터는

2016년 기준 코스피 및 코스닥 시장에 상장된 1,772개 기업에 대한 신용등급 데이터

- ◆ 독립변수는 총 15개이며 [표 8-3]에서 제시한 신용등급 관련 변수를 대부분 이용함

- ◆ 래피드마이너에서는 속성을 정의해야 함.

속성 정의에는 속성 이름(attribute name), 속성 값 유형(attribute value type), 속성 역할(attribute role) 등이 정의되어야 함

8.2 기업신용등급예측모형 구축 사례

(1) 신용등급데이터 및 연구변수

데이터 유형	래피드마이너 속성 값 유형
범주 유형	nominal
숫자 유형	numeric
정수 유형	integer
실수 유형	real
텍스트 유형	text
이항 범주 유형	binominal
다항 범주 유형	polynominal
날짜 시간 유형	date time
날짜 유형	date
시간 유형	time

[기타] 래피드마이너의 속성 값 유형

8.2 기업신용등급예측모형 구축 사례

(1) 신용등급데이터 및 연구변수

변수명	값 유형	역할	설명
기업명	text	id	코스피 및 코스닥 상장기업
신용평점	integer	사용안함	1(최우수) ~ 10(최하)
신용등급(종속변수)	binominal	label	0(우량기업)/1(불량기업)
유동비율(X_1)	real	attribute	유동성비율
당좌비율(X_2)	real	attribute	
부채비율(X_3)	real	attribute	레버리지비율(안정성)
영업이익이자보상비율(X_4)	real	attribute	
차입금/자기자본(X_5)	real	attribute	
차입금/매출액(X_6)	real	attribute	

[표 8-3] 신용등급예측을 위한 연구변수와 속성 정의(1)

8.2 기업신용등급예측모형 구축 사례

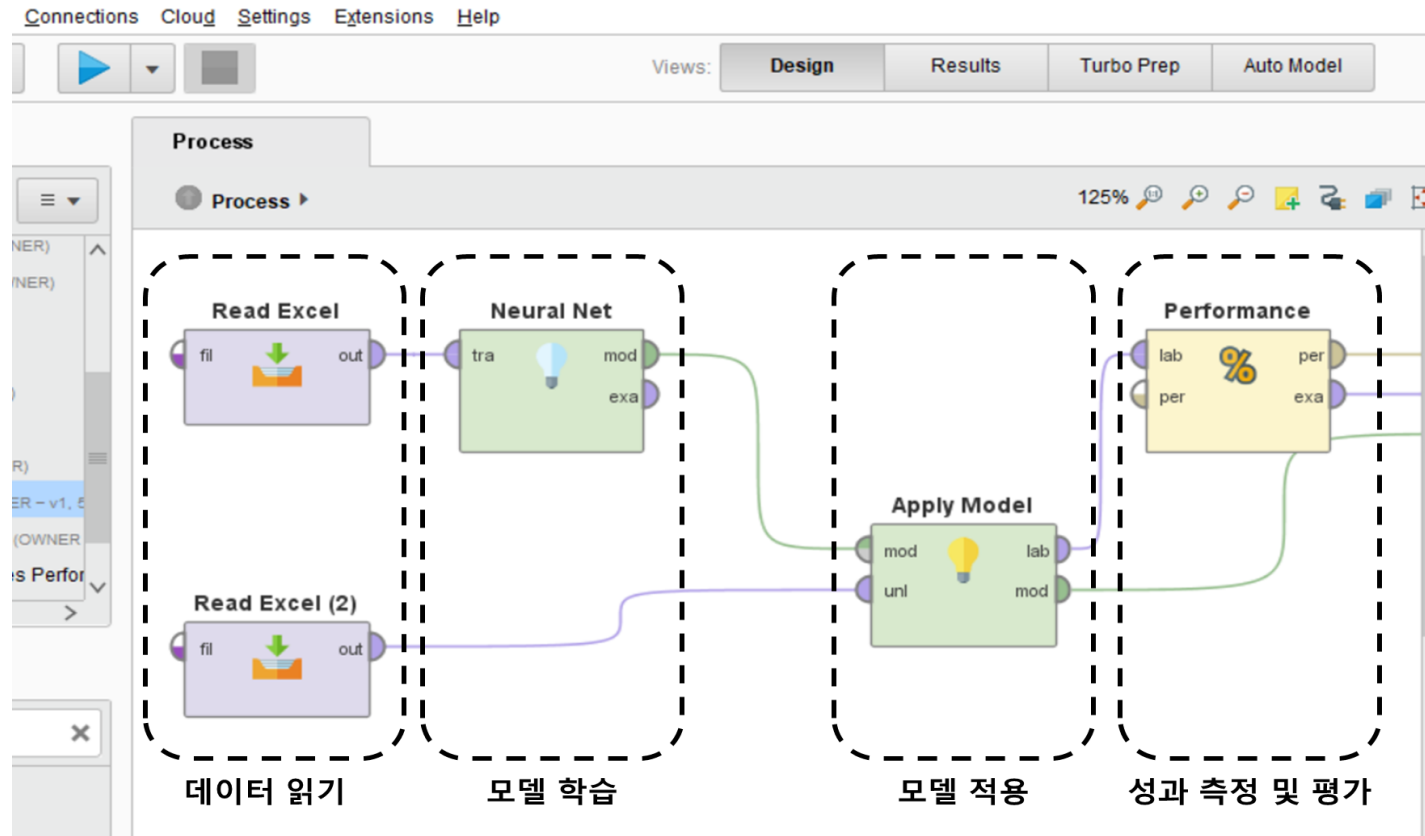
(1) 신용등급데이터 및 연구변수

변수명	값 유형	역할	설명
총자산증가율(X_7)	real	attribute	성장성비율
매출액증가율(X_8)	real	attribute	
순이익증가율(X_9)	real	attribute	
재고자산회전율(X_{10})	real	attribute	활동성비율
매출채권회전율(X_{11})	real	attribute	
총자본회전율(X_{12})	real	attribute	
총자본 순이익률(X_{13})	real	attribute	수익성비율
매출액순이익률(X_{14})	real	attribute	
자기자본순이익률(X_{15})	real	attribute	

[표 8-3] 신용등급예측을 위한 연구변수와 속성 정의(2)

8.2 기업신용등급예측모형 구축 사례

(2) 신경망을 이용한 기업신용등급예측모형



[그림 8-1] 신경망을 이용한 기업신용등급예측모형

8.2 기업신용등급예측모형 구축 사례

(2) 신경망을 이용한 기업신용등급예측모형

accuracy(정확도): 87.15% (은닉층: 2, 은닉 노드의 수: 7)

	true 1 (실제 불량기업)	true 0 (실제 우량기업)	예측정확도
pred. 1 (예측 불량기업)	222개	51개	81.32%
pred. 0 (예측 우량기업)	40개	395개	90.80%
class recall	84.73%	88.57%	—

[표 8-5] 신경망 기반 신용등급예측모형의 예측성과

핀테크와 스마트 금융

8장 / 11주차 2차시

8.3

P2P 대출 개인채무불이행 예측모형 구축사례: P2P금융과 P2P대출

◆ P2P 금융

- 금융기관을 통하지 않고 인터넷·모바일 등으로 개인 간에 직접적으로 금융거래가 이루어지는 것을 말함

◆ P2P 대출

- P2P 금융의 대표적인 서비스로 금융기관의 중개 없이 온라인상에서 차입자가 자금이 필요한 사정 및 재무상황에 관한 정보를 게시하면, 투자자들이 이를 참고하여 자금을 빌려주는 거래를 말함

- P2P 대출은 대출자(차입자)들이 온라인 P2P 플랫폼을 통해 대출 신청하면 P2P 플랫폼 기업이 심사하고, 이를 공개하여 불특정 다수가 여유자금을 빌려주고 이자를 받는 대출중개서비스를 말함

◆ 온라인을 통해 모든 대출과정을 자동화하여 지점운영비용, 인건비, 대출영업비용 등의 경비 지출을 최소화하여 대출자에게는 보다 낮은 금리를, 투자자에게는 보다 높은 수익을 제공하고 있음

8.3 P2P 대출거래의 구성요소

◆ P2P 대출서비스

- 담보물 필요 없이 대출경매(주로 역경매)를 통해 투자자들이 대출을 심사하여 결정함

◆ P2P 대출

- 일반 금융기관 대출과 달리 온라인 플랫폼 상에서 차용인의 정보만을 가지고 투자를 결정하기 때문에 사회관계망(SNS)을 통한 차용인과 투자자 간 상호작용이 매우 중요한 요소로 작용함
- 차용인의 인구통계정보, 개인신용정보, 대출정보, 차용인의 대출을 상환하기 위한 계획 및 노력 등의 주관적·비정형 정보가 대출결정의 중요한 요소로 작용함

◆ P2P 대출거래의 주체는 아래의 다섯가지로 구성됨.

- 차용인
- P2P 대출플랫폼
- 투자자
- 이사회
- 감독기관

8.3 P2P 대출거래의 구성요소

◆ 차용인과 투자자

- P2P 대출업체를 통해 본인에게 필요한 대출유형서비스를 선택하고, 대출승인과 대출금리에 관한 경매에 참여함

◆ P2P 대출플랫폼

- 중개자의 역할로서 차용인으로부터 매달 원금과 이자를 받고 이를 투자자에게 돌려줌

◆ 이사회

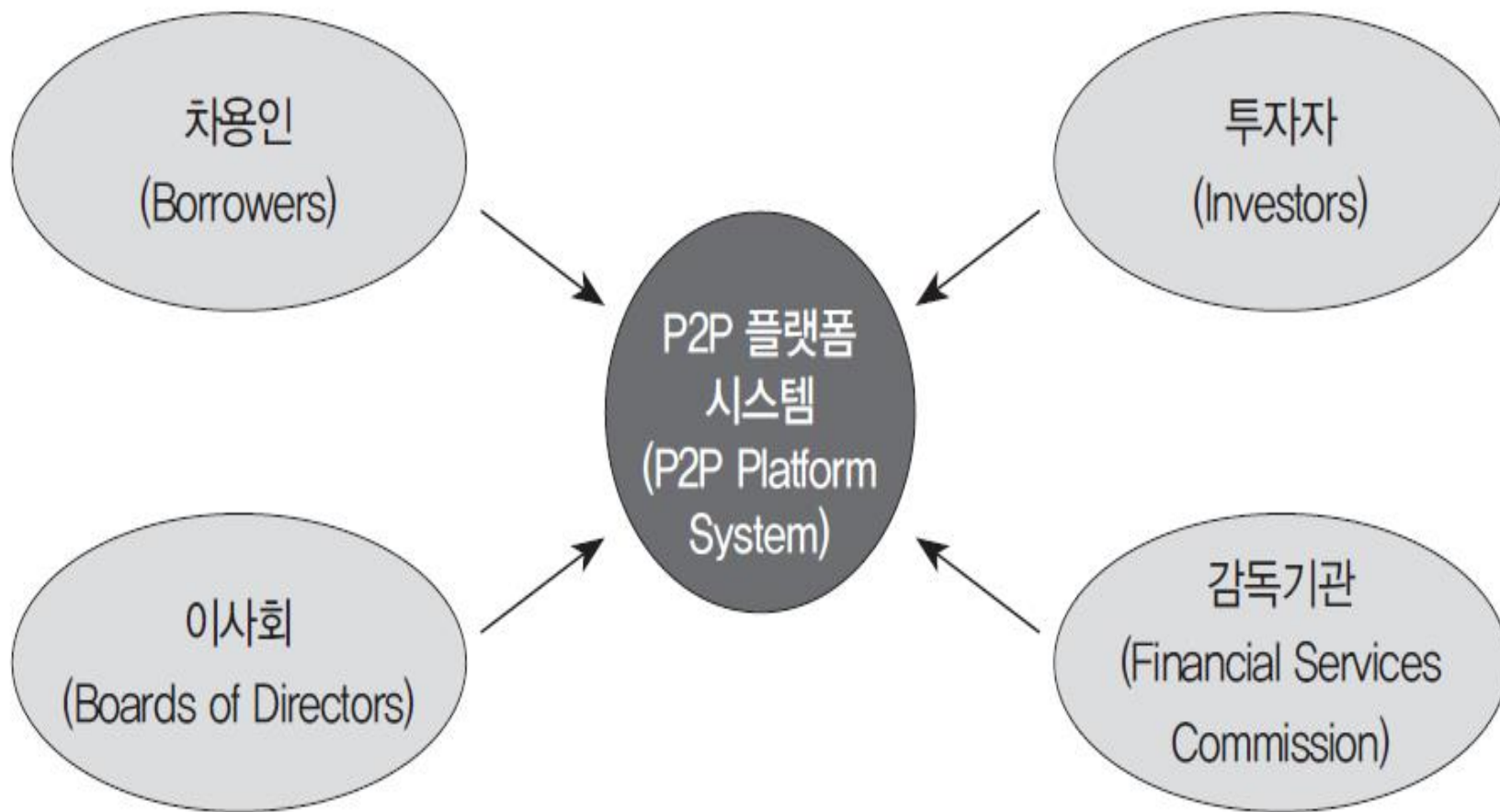
- P2P이용자에게 향상된 금융서비스를 제공하고자 플랫폼의 안정성 관리에 노력을 기울임
(새로운 비즈니스 모델과 금융상품 개발)

◆ 감독당국

- P2P 대출업체의 재무건전성 검토 및 부실플랫폼 퇴출, P2P 대출거래의 비합리적 규제완화, 금융소비자보호와 신뢰성 확보 방안을 마련함

8.3

P2P 대출거래의 주체



8.3 P2P 대출거래 (3단계)

◆ 1단계

- 차용인(대출자)들이 본인의 재무상태와 대출목적, 대출금액을 P2P 플랫폼에 등록하면 투자자들은 산출된 정보들과 대출자의 신용정보를 바탕으로 대출 결정 내림

◆ 2단계

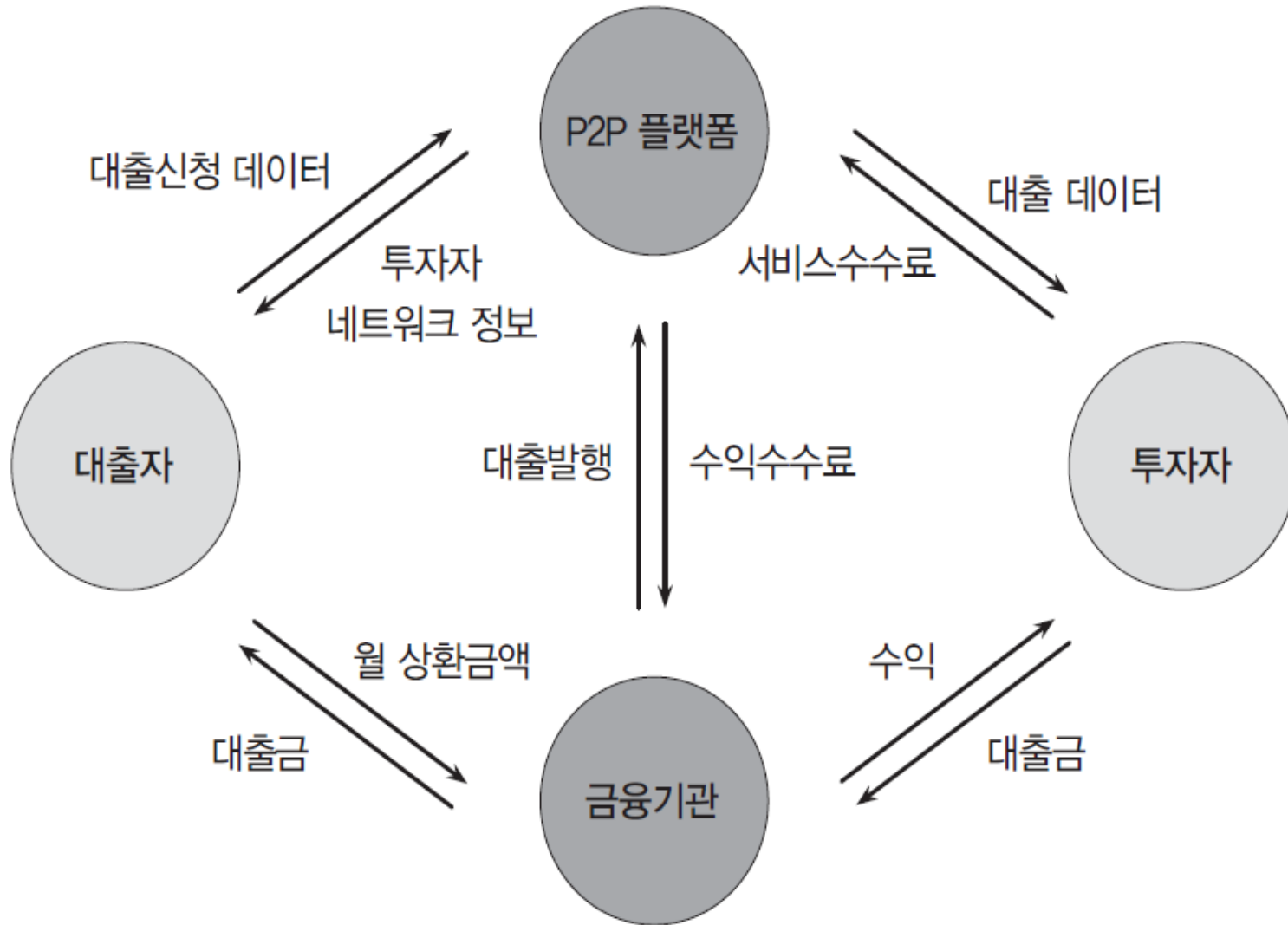
- 투자자들은 투자금액으로 대출자의 대출금액을 입찰하고, 대출자는 최대 이자 한도 내의 이자를 입찰함
- 낮은 이자를 제시한 투자자가 입찰에 성공하고 대출자에게 자금을 대출할 수 있음

◆ 3단계

- 경매 종료되면 P2P 대출업체의 제휴 저축은행이 투자자 계좌에서 대출액을 출금하고 이를 저축은행에 대출보증으로 입금함
- 대출자는 대출승인비 등의 비용(수수료)을 차감하고 저축은행에서 대출받음
- 대출자가 대출상환에 성공하면 저축은행은 신용등급 및 신용점수에 반영함

8.3

P2P 대출거래의 구조



8.3 P2P 대출거래 공개용 데이터베이스

(1) 렌딩클럽(Lending Club)의 P2P 대출거래 데이터베이스

◆ 렌딩클럽(Lending Club)

- 2006년에 설립됨
- 미국 최대의 P2P 금융기업으로 2014년에 P2P 기업 최초로 뉴욕증권거래소에 상장됨
- 금융빅데이터 분석, 기계학습을 포함한 데이터마이닝 기법 등의 ICT 역량을 활용하여 다양한 대출상품과 수익 모델을 개발함
- 채무상태와 채무정보, 발행된 대출거래 데이터, 렌딩클럽의 신용정책에 충족하지 못한 거절된 대출신청의 세부사항에 대한 데이터를 공개하고 있음
- 2007년부터 현재까지 약 250만 개의 P2P 대출거래 데이터(총 151개 변수)를 홈페이지에서 무료로 제공하고 있음

8.3

P2P 대출거래 공개용 데이터베이스

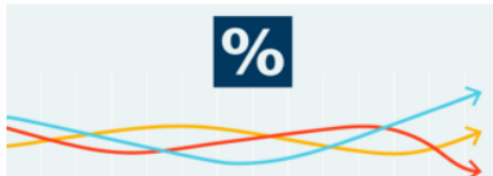
(1)

렌딩클럽(Lending Club)의 P2P 대출거래 데이터베이스

[BORROW ▾](#)[INVEST ▾](#)[ACCOUNT](#)[SETTINGS](#)[HELP](#)[SIGN OUT](#)

Loan Statistics

Want to slice and dice the data? Below find a variety of additional data on credit performance, loans, payment history, and summary snapshots.



Historical Loan Issuance Data

2020 Q1 ▾

Download

These files contain loan data for all loans issued through the time period stated, including the current loan status (current, late, fully paid, etc.) and latest payment information.

Data Dictionary

Includes definitions for data attributes included in the historical loan issuance data file above.



Summary Data Files

[Cumulative Charge-Off Rates & Projected Loss Curves](#)

[Delinquency Rates](#)

[Prepayment Rates](#)

[Monthly Snapshot](#)

[Sample Quarterly Report](#)

Actual results may vary.



Payment History Files

[Payments Made to Investors and LendingClub](#)

[Payments Made to Investors](#)

Due to the large file size, statistical software is needed to download in full.

8.3 P2P 대출거래 공개용 데이터베이스

(2) 렌딩클럽(Lending Club)의 P2P 대출상환결정 주요 변수

◆ 렌딩클럽의 P2P 대출거래 데이터베이스의 주요 변수

- 인구통계학적 특성 및 차입자정보, 대출정보, 신용정보 등

(1) 인구통계적 특성 및 차입자정보

- 나이, 성별, 주택소유형태, 혼인여부, 고용상태, 근무경력, 연소득, 예금평잔액, 총부채상환비율 등으로 구성됨

(2) 대출정보

- 차입자의 대출금액, 대출목적, 대출이자율, 대출횟수 등으로 구성됨

(3) 신용정보

- 차입자의 신용등급, 신용파산경험, 정보공개횟수, 신용조회횟수, 연체횟수, 모기지(mortgage) 대출계좌수, 신용카드 계좌 수 등으로 구성됨

8.3 P2P 대출거래 공개용 데이터베이스

분류	변수	설명
인구 통계적 특성 및 차입자 정보 (9개)	나이	연령대를 10대로 구분
	성별	남녀 구분
	주택소유형태	자가, 전세, 월세, 기타
	혼인 여부	결혼 여부
	고용 상태	상근직, 비상근직, 자영업, 무직 등
	근무 경력	차입자의 근무 경력(월)
	연소득	연간 소득액
	예금 평잔액	6개월 평균 잔고액
	총부채상환비율	총부채/소득
대출 정보 (4개)	대출 금액	대출 신청 금액 로그(log)화
	대출 목적	생계필수비용(주거비, 생활비, 의료비), 사업용, 부채 상환용
	대출 이자율	온라인 P2P 차입자의 대출 이자율
	대출 횟수	기존에 대출받은 횟수

[표 8-6] 렌딩클럽(Lending Club)의 P2P 대출상환결정 주요 변수(1)

8.3 P2P 대출거래 공개용 데이터베이스

분류	변수	설명
신용 정보 (7개)	신용 등급	차입자의 신용등급
	신용 파산 경험	파산, 면책, 워크아웃, 회생 등의 경험
	정보 공개 횟수	차입자의 공개된 개인 신상 정보 공개 횟수
	신용 조회 횟수	차입자의 신용 조회 횟수
	연체 횟수	차입자의 최근 2년간 연체 횟수
	모기지(mortgage) 계좌 수	모기지론(mortgage loan) 계좌 수
	신용카드 계좌 수	차입자가 발급받은 신용카드 계좌 수
사회적 교류 활동 (5개)	질의응답 수	차입자, 투자자의 P2P 질의응답 게시판 질의응답 횟수
	대출 시도 횟수	대출 성공까지의 경매 시도 횟수
	대출 상환 투표율 (긍정, 부정)	대출 상환 성공 가능성 여부에 대한 대출 전 투표 실시
	대출 신청 사연	동정, 도움, 긴급, 성실성 등 차입자의 언어 표현 정보
	대출 상환 계획	차입자가 구체적인 대출 상환 계획을 제시한 경우 1의 값을 갖는 더미변수

[표 8-6] 렌딩클럽(Lending Club)의 P2P 대출상환결정 주요 변수(2)

8.3 P2P 대출거래의 채무불이행예측모형 구축

(1) 자료수집 및 변수선정

- ◆ P2P 대출의 채무불이행예측모형 구축을 위한 실습용 데이터는 렌딩클럽(Lending Club)의 공개용 데이터베이스를 활용함
- ◆ 2016년 1월부터 2017년 12월까지 렌딩클럽의 P2P 대출거래 데이터를 수집함, 최종적으로 상환완료(Fully Paid)와 채무불이행(Default) 속성만을 종속변수로 활용함
- ◆ 본 실습에서 사용된 P2P 대출거래 데이터 수는 143,639건이며, 이 중 상환완료된 대출은 113,670건이고, 채무불이행은 29,969건
- ◆ 인구통계학적 변수를 포함한 차입자정보 관련 변수, 대출정보변수, 재무 및 신용변수로 이루어진 13개의 독립변수를 선정함

8.3

P2P 대출거래의 채무불이행예측모형 구축

변수명	값 유형	역할	설명
대출 상환 상태(종속변수)	binominal	label	0(상환 완료) /1(채무불이행)
주택소유형태(X_1)	polynominal	attribute	자가, 전세, 월세, 기타
근무경력(X_2)	integer	attribute	차입자의 근무경력(월)
연간 소득액(X_3)	real	attribute	연간 소득액
총부채 상환비율(X_4)	real	attribute	총부채/소득
대출 금액(X_5)	integer	attribute	대출 신청 금액의 로그(log)화
대출 목적(X_6)	polynominal	attribute	생계 필수 비용, 차량 구매, 사업용, 부채 상환용
대출 이자율(X_7)	real	attribute	차입자의 대출 이자율

[표 8-8] P2P 대출거래시스템의 대출상환결정 주요 변수(1)

8.3 P2P 대출거래의 채무불이행예측모형 구축

변수명	값 유형	역할	설명
신용등급(X_8)	polynominal	attribute	렌딩클럽이 부여한 차입자의 신용등급
신용 파산 기록(X_9)	integer	attribute	차입자의 파산, 면책, 워크아웃, 회생 등의 기록 유무
연체 횟수(X_{10})	integer	attribute	지난 2년간 연체 횟수
금융거래 건수(X_{11})	integer	attribute	금융거래 횟수
신용카드 구매가능금액(X_{12})	integer	attribute	신용카드 최대 구매가능금액
모기지 대출 계좌 수(X_{13})	integer	attribute	모기지론(mortgage loan) 계좌 수
신용카드 계좌 수(X_{14})	integer	attribute	차입자가 발급받은 신용카드 계좌 수

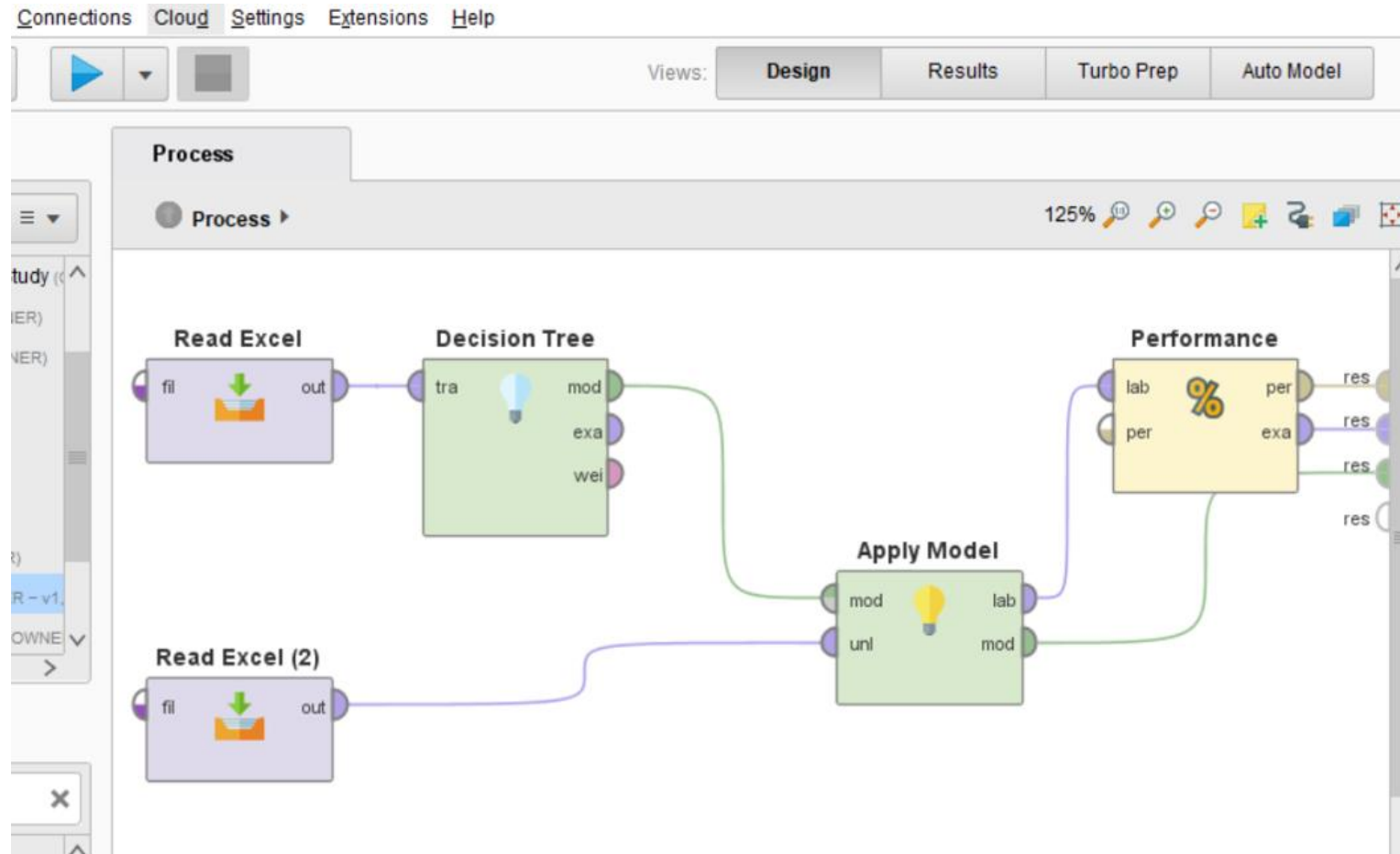
[표 8-8] P2P 대출거래시스템의 대출상환결정 주요 변수(2)

8.3

P2P 대출거래의 채무불이행예측모형 구축

(2)

P2P 대출거래의 채무불이행예측모형 구축: 의사결정나무



[그림 8-5] 의사결정나무(C5.0)을 이용한 P2P 대출거래의 채무불이행예측모형

8.4 빅데이터와 인공지능: 인공지능의 개념

◆ 인공지능(AI)

- 인간의 학습 능력, 추론 능력, 지각 능력, 자연어 이해 능력 등을 컴퓨터 프로그램으로 실현한 기술 - - 컴퓨터과 학자 존 매카시(John McCarthy)가 1956년 다트머스 콘퍼런스에서 사용한 표현
- 오늘날에는 기억, 지각, 이해, 학습, 연상, 추론 등 인간의 지성을 필요로 하는 행위를 기계를 통해 실현하고자 하는 학문 또는 기술의 총칭

◆ 마이크로소프트의 공동창업자인 폴 앨런(Paul Allen)

- 현존하는 세상의 모든 지식(인터넷에 축적된 빅데이터를 포함)을 인공지능 시스템과 연결하는 슈퍼 인공지능 프로그램인 **디지털 아리스토텔레스(아리스토, Aristo)** 개발함

◆ 아리스토는 방대한 지식을 컴퓨터가 처리 가능한 형태로 변환하여

저장한 지식통합 시스템을 구축하고, 자연어처리 기술을 통해 인간의 질문을 이해하고, 다양한 수준의 질문에 답을 제공함

8.4 인공지능에서 빅데이터가 중요한 이유

(1) 대량의 데이터와 반복적인 작업을 빠르게 처리할 수 있는 인공지능 기술은
데이터에 존재하는 패턴이나 특징을 분석하고 스스로 자동학습을 하도록 지원함

(2)

▶ 인공지능

- 점진적인 학습 알고리즘을 통해 스스로를 개선하고 데이터가 프로그래밍을 수행하도록 지원함

▶ 딥러닝(deep learning, 심층학습)

- 컴퓨팅 파워의 발전과 학습 기법의 개선을 바탕으로 심층신경망(Deep Neural Network)을 활용함

- 데이터의 구조와 규칙성을 찾아내고 스스로 학습하여 최적화 관련 의사결정을 지원함

(3) 심층신경망 분석을 통해 다양한 금융의사결정 문제에 관한 향상된 정확도를 제공함

다양한 사용자들이 인공지능 플랫폼을 이용할수록 더 높은 정확도로 원하는 결과 획득 가능

8.4 인공지능에서 빅데이터가 중요한 이유

(4) 인공지능

- 데이터 활용도를 극대화하며 빅데이터 보유는 차별화된 경쟁우위 요소로 작용함
- 알고리즘이 자체 학습을 수행하면 데이터 자체가 지적재산이 될 수 있음
- 경쟁이 치열한 산업에서는 유사한 기술을 보유하여 경쟁우위를 유지하는 기업보다 최상의 데이터(품질이 보증된 빅데이터)를 보유한 기업이 살아남게 됨

◆ 인공지능과 빅데이터 분석 기술을 성공적으로 연계하기 위해서

인공지능 기반 데이터(AI-ready data) 획득 및 빅데이터 관리 시스템 도입이 필요함

- 인공지능 시스템 구동을 위한 필수 요건:

인공지능에 최적화된 충분한 양의 정확한 데이터(빅데이터 센터 등 인프라 중요함)

8.4 금융빅데이터와 인공지능

◆ 금융빅데이터

- 다른 산업 데이터보다 정형화(structured)되어 있어서 데이터 품질이 높음
- 데이터 거래소를 통한 금융빅데이터의 구축은 인공지능 산업의 성장에 기여할 수 있음

◆ 최근 금융 업계

- 이자수익 감소, 규제강화, 보안위협 증가, 고객 요구사항 증가,
비대면 서비스 증가에 따른 비용 부담 등은 물론 핀테크 스타트업과의 경쟁도 치열해져
인공지능 기술의 도입을 통한 탈금융화 및 디지털 전환의 필요성 증가

◆ 인공지능을 활용한 금융서비스

- 주식, 채권, 외환 등에 대한 투자의사결정뿐 아니라 대출 신청 승인 여부,
효과적 자산배분, 금융상담, 핵심의사결정까지 인간의 영역을 빠르게 잠식하고 있음

8.4

빅데이터와 인공지능 융합 비즈니스: 로보어드바이저

◆ 로보어드바이저 서비스

- 금융산업에서 대표적인 빅데이터와 인공지능 융합 비즈니스 모델은
- 금융빅데이터, 머신러닝(기계학습), 딥러닝(심층학습) 기반의 알고리즘으로 구현되는 인공지능 기술
- 사전에 온라인으로 작성된 투자자(금융소비자)의 정보를 활용하여 투자 및 위험성향을 분석해 포트폴리오(portfolio)를 구성하고, 이를 바탕으로 맞춤형 투자자문과 자산운용을 수행함
- 로보어드바이저의 포트폴리오 배분 프로세스 단계(5단계)

8.4

로보어드바이저의 포트폴리오 배분 프로세스 단계

- (1) 개별 질문을 통해 투자자의 투자성향과 위험성향을 파악하고 투자목적을 분석하여
(customer profiling) 투자자금의 성격 파악
- (2) 투자자의 성향 및 목적에 따른 자산군별 투자비중을 결정함
(asset allocation, 자산배분)
- (3) 금융빅데이터 기반으로 인공지능 알고리즘을 이용하여
최적의 맞춤 포트폴리오를 추천 선택하여 자산군별 최적의 금융상품을 추천함
(portfolio selection, 포트폴리오 추천)
- (4) 로보어드바이저가 추천한 투자를 집행함 (trade execution)
- (5) 투자가 실행되는 시장과 투자자산을 모니터링(monitoring)하면서
포트폴리오를 자동으로 재조정(portfolio rebalancing)

8.4

빅데이터와 인공지능 융합 비즈니스: 인슈어테크

◆ 인슈어테크(InsurTech)

- 보험(Insurance)과 기술(Technology)의 합성어로

기존의 보험 서비스에 4차 산업혁명 기술(빅데이터 분석, 인공지능)을 융합한 보험 서비스의 혁신

◆ 보험 핀테크

- 빅데이터 분석, 인공지능, 블록체인 등 활용해 보험금 청구 간소화 서비스,

빅데이터 분석을 통한 고객 맞춤형 보험 등과 같은 기존과 다른 보험 서비스를 제공한다

◆ 전체 보험 가입자에게 동일하게 적용하던 보험요율을 빅데이터 분석을 통해 다르게 적용할 수 있음

◆ 보험상담 업무도 로봇이 대행할 수 있고, 보험빅데이터 관리를 통한 효과적인 영업을 하며,
블록체인을 이용해 안전한 결제 시스템 구축

8.4 인슈어테크로 인한 보험산업의 가치사슬 변화

◆ 보험산업의 가치사슬 분야: 상품개발, 유통 및 판매, 보험가입심사, 보험금 지급관리, 마케팅 및 고객관리 등

가치사슬 단계	세부 내용
상품개발	위험요율 산출 기법이 다양화되고 보험빅데이터를 활용한 고객 맞춤형 상품개발 가능
유통/판매	온라인, 모바일 등 다양한 유통채널을 통한 신규 고객 확보 용이
보험가입심사	자동화된 언더라이팅 시스템을 통해 처리 시간 단축 및 정확도 향상, 비용 효율성 높아짐
보험금 지급관리	보험금 청구절차를 간소화하고, 보험사기를 사전에 방지할 가능성 높아짐
마케팅 및 고객관리	온라인, 모바일에 최적화된 사이트를 통해 고객을 확보하고, 자동화 시스템을 통해 개별 고객과의 커뮤니케이션 활성화 가능

8.5 8장 핵심정리

1. 블록체인(Blockchain)의 정의

- 분산형 데이터베이스와 유사한 형태로 모든 구성원이 네트워크를 통해 데이터를 검증 및 저장하여 특정인의 임의적인 조작이 어렵게 설계된 저장 플랫폼

2. 기업신용등급예측 모형 구축 사례

- 신용등급의 정의: 채무자의 채무불이행 가능성에 대해 기업의 경영 능력, 운영 능력, 시장에서의 경쟁력 등 다양한 요인이 반영된 미래 현금흐름과 재무적 융통성을 고려하여 일정한 기호로 나타낸 것
- 회사채 신용등급예측: 대출 또는 투자 대상 기업의 신용등급 수준을 판별하는 신용등급평가 문제
- 기업부도예측: 대출 또는 투자 대상 기업의 부실 여부를 판별

8.5 8장 핵심정리

3. P2P 대출 개인채무불이행 예측 모형 구축 사례

▶ P2P 대출

- 대출자(차입자)들이 온라인 P2P 플랫폼을 통해 대출을 신청하면, P2P 플랫폼 기업이 심사하고 결과를 공개하여 불특정 다수가 여유자금을 빌려주고 이자를 받는 대출중개 서비스

▶ 금융규제 샌드박스(Regulatory Sandbox)

- 혁신적인 P2P 금융상품을 규제의 제약 없이 일정 기간 테스트해볼 기회를 제공하는 것

4. 빅데이터와 인공지능 개요

▶ 인공지능의 개념

- 인간의 학습 능력, 추론 능력, 지각 능력, 자연어 이해 능력 등을 컴퓨터 프로그램으로 실현한 기술

8.5 8장 핵심정리

4. 빅데이터와 인공지능 개요

▶ 빅데이터와 인공지능 연계

- 인공지능이 클라우드 서버의 빅데이터를 분석하여 의사결정문제에 대한 최적화(예측)함으로써
가치 창출 (데이터 수집 → 저장 및 분석 → 가치 창출 → 최적화 과정)

◆ 빅데이터와 인공지능 융합 비즈니스, **로보어드바이저**

▶ 로보어드바이저

- 사전에 온라인으로 작성된 투자자(금융소비자)의 정보를 활용하여
투자 및 위험성향을 분석해 포트폴리오를 구성하고,
이를 바탕으로 맞춤형 투자자문과 자산운용을 수행함

◆ 빅데이터와 인공지능 융합 비즈니스, **인슈어테크(InsurTech)**

- 인슈어테크: 보험(Insurance)과 기술(Technology)의 합성어로,
기존의 보험서비스에 4차 산업혁명 기술(빅데이터 분석, 인공지능)을 융합한 보험서비스의 혁신