

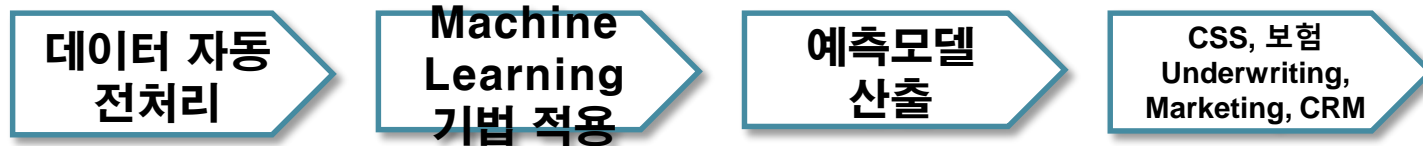


# Solidware

We turn data into money.

## Machine Learning을 활용한 금융 데이터 분석

# 솔리드웨어 프로젝트 현황



자동차 보험 고객 손해액  
예측 솔루션 개발



개인 대출고객 연체율  
예측 솔루션 개발



Machine Learning 기반  
대안적 U/W 모형 개발



소매 중금리  
신용평가모형 개발



AS 신용평가 모형  
고도화 프로젝트

**“기업을 위한 실질적 가치 창출용  
인공지능 솔루션”**

핵심 기술 :

**Machine Learning**

가장 발전된 빅 데이터 분석 기술

컴퓨터 과학 및 인공지능의 하위 분야로서,

데이터를 통한 알고리즘 트레이닝에 기초하며,

복잡하고 정교한 예측 함수를 도출하여  
추정하고자 하는 미래의 값들을 구합니다.



# Machine Learning은 다양한 IT 기업에서 활용된 바 있으며, 수많은 실적을 통해 이미 성능이 검증된 기술입니다.

## Machine Learning 성공 사례



광고업, 검색 엔진, 스팸 탐지, 음성 인식, ...



광고업, 검색 엔진, Kinect, ...



광고업, 상품 추천, 상품 인식, ...



광고업, 뉴스피드, 친구 추천, 얼굴 인식, ...

# 최근 Machine Learning이 부각되는 이유는?

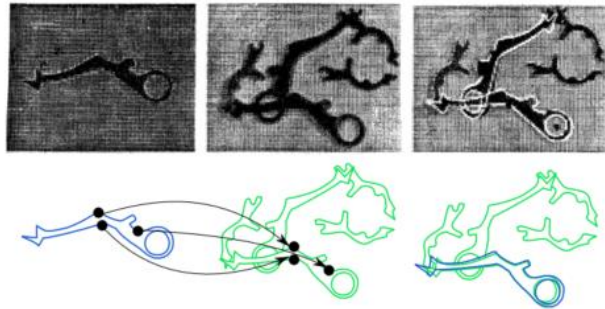


- 
1. 고성능 컴퓨터의 보편화
  2. 보다 향상된 알고리즘
  3. 증가된 데이터의 양

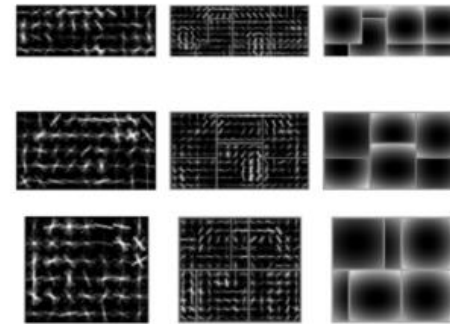
“현재 존재하는 **모든 데이터의 90%**를 지난 **2년 간** 생산했을 정도로 우리는 수많은 데이터를 생산하고 있다 ”

- Petter Bae Brandtzæg, SINTEF ICT

# ‘Computer Vision’ 분야 내 Machine Learning 연혁



**1980s**  
정렬법

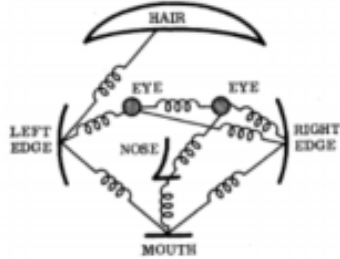


**2000s**  
가변형 모델

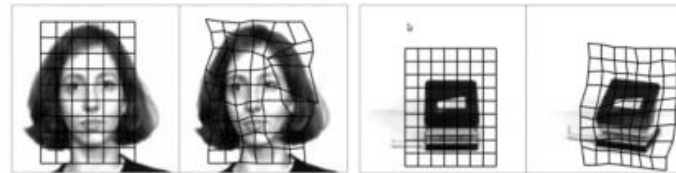
모델기반

통합

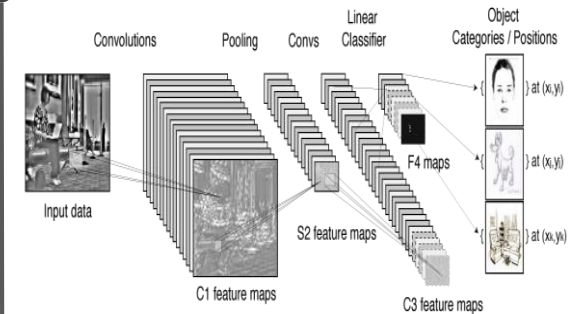
데이터기반



**1970s**  
수공모델



**1990s**  
격자모델



**2010s**  
융합네트워크

# Computer Vision 연구원이 사용하는 데이터의 양은?

2004  
Caltech 101  
10K 이미지 사용

2005-2010  
Pascal VOC  
2K → 30K 물체 사용

2010-2015  
Image Net  
10M → 15M 이미지 사용



Image source: <http://www.vision.caltech.edu/>

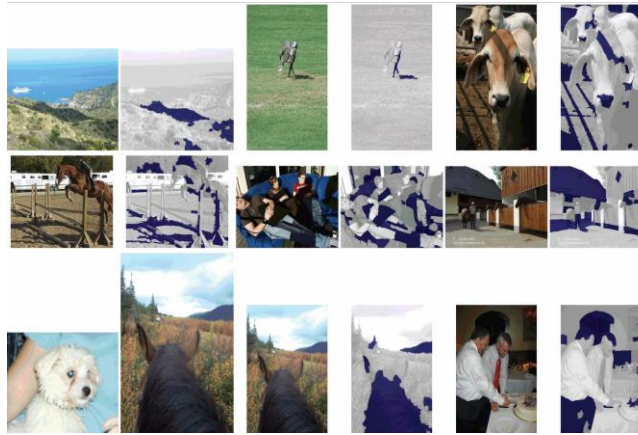
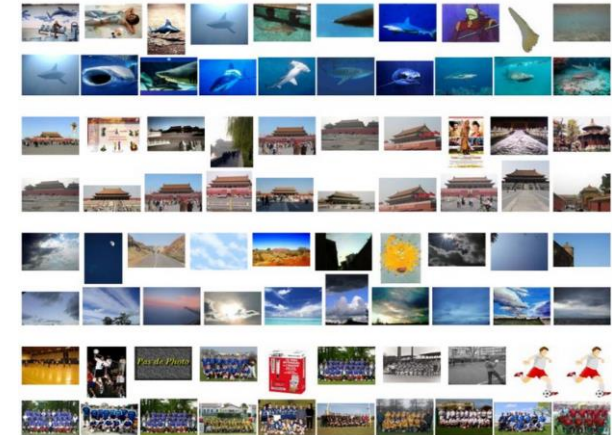


Image source: <http://doi.ieeecomputersociety.org/>



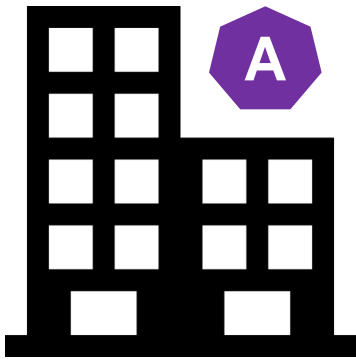
<http://www.image-net.org/>



Image source - [Smartdatacollective.com](https://www.smartdatacollective.com)

# 데이터의 양 만큼 중요한 알고리즘의 역할

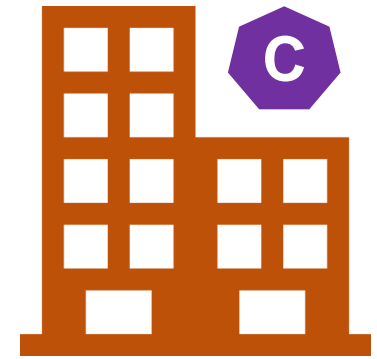
해결하고자 하는 문제와 회사 목적의 특징에 맞춘 알고리즘 필요



“딥러닝”



“신경망”



“로지스틱 회귀분석”

Deep learning

Boosting

K-Nearest neighbors

Artificial neural networks

Support vector machines

Bayesian networks

Sparse dictionary learning

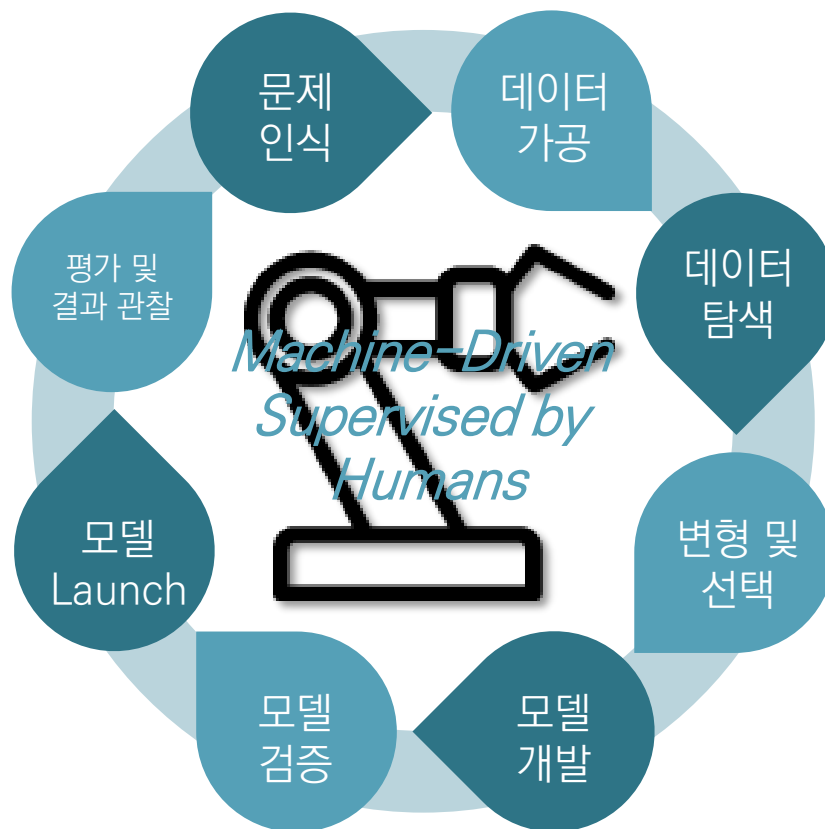
Regression forest

# 머신러닝 일반 방법론

- ◆ 사람이 수행하는 통계 기법이 아닌 자동화된 Machine-Driven 분석 기법을 통해 전략 최적화 및 운영 효율성 극대화



- 통계 기반 분석 프로세스의 모든 단계 사용자 인풋 필요
- 사람의 직관 및 배경 지식이 분석 방법 및 결과 판단 기준의 척도로 오판단 가능성 존재
- 기법 특성상 제한된 양의 데이터 활용



- Machine Learning 알고리즘을 통한 분석 프로세스의 **자동화**로 효율적이고 더 정교한 데이터 분석이 가능함
- **사람의 개입이 최소화**되어 오판단 가능성 최소화
- 기법 특성상 모든 데이터 활용 가능

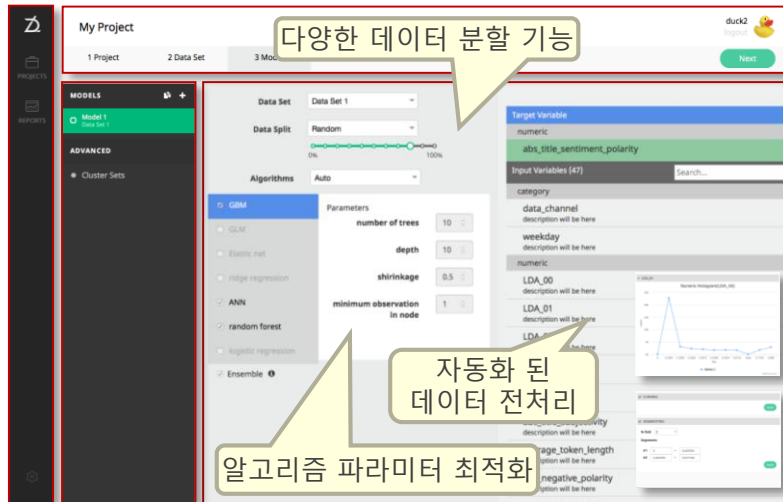
# **Machine Learning for Everyone**

## **: How to internalize the tech**

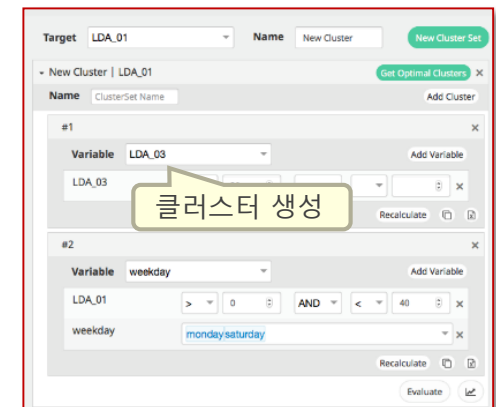
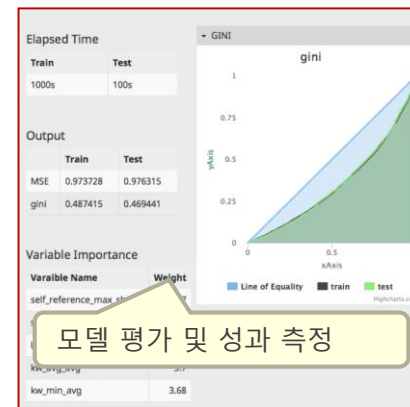
# DAVinCI LABS

## DAVinCI LABS는

- ✓ 머신 러닝 또는 컴퓨터 공학을 전공하지 않은 **비전문 사용자**도 그 가치를 충분히 누릴 수 있도록 설계된 세계 유일의 Machine Learning 인공지능 데이터 분석 솔루션
- ✓ 실제로 다양한 기업 프로젝트 수행을 통해 축적된 **솔리드웨어 팀 노하우가 그대로 담겨져 있는 기업형 실용 기능**을 제공, 비즈니스 실무진의 편의성을 최대화
  - 과적합 등의 예측 오류 최소화
  - 설명력 저하(black box)를 막기 위한 풍부한 시각화 및 리포팅
- ✓ 데이터의 가공부터 리포트까지, **end-to-end의 seamless한 생태계**를 구축
- ✓ 기존 **빅데이터/Machine Learning**의 한계를 극복한 최적의 예측 솔루션
  - 과적합 등의 예측 오류 최소화
  - 설명력 저하(black box)를 막기 위한 풍부한 시각화 및 리포팅







3. 자동화 된 스마트 데이터 전처리 지원 (결측값 / 중복값 / 극단값 / 변수타입 등)
4. 알고리즘 자동최적화를 통한 예측력 향상 및 과적합 방지
5. 분석 전과정에 걸친 뛰어난 시각화
6. 다양한 분석 결과를 비교, 출력 할 수 있는 리포팅 시스템



1. 편리한 분석 환경 – 자체 고안된 직관적 UI/UX 지원
2. 실무 사용자의 니즈 반영
  - ✓ 비즈니스 룰 적용(클러스터 생성) 등 다양한 기능
  - ✓ 도출된 예측함수의 추출/통합 지원을 통한 활용성 강화



# DAVinCI LABS 분석 프로세스별 모듈 ILLUSTRATION

전처리 모듈	분석 모듈	시각화 모듈	리포팅 모듈
 <p><b>자동 전처리 기능</b></p>	 <p><b>자동적으로 최적 알고리즘 선택</b></p>	 <p><b>직관적인 시각화를 통한 분석 결과 표기</b></p>	 <p><b>다양한 알고리즘 적용결과간 비교 분석</b></p>
<ul style="list-style-type: none"> <li>▪ 데이터를 자동으로 머신 러닝 적용 가능 데이터로 변환</li> <li>▪ 각종 결측값/극단값 및 중복값 등에 robust한 스마트 데이터 전처리 제공</li> <li>▪ 변수타입 자동 설정</li> <li>▪ 최적 변수변환 지원</li> <li>▪ 중요항목 자동분석 및 feature engineering 수행을 통한 과적합 최소화</li> </ul>	<ul style="list-style-type: none"> <li>▪ 인공지능 기반의 알고리즘 파라미터 세팅 최적화</li> <li>▪ 알고리즘들의 선형 조합으로 예측력 극대화</li> <li>▪ 최적 군집/클러스터 자동 생성 기능</li> <li>▪ 시계열/random 등 다양한 데이터 분할 기능 제공</li> <li>▪ 알고리즘간 조합 가능</li> </ul>	<ul style="list-style-type: none"> <li>▪ 분석 프로세스 전반에 걸친 각 파트별 시각화</li> <li>▪ 변수 및 군집 분포도 EDA를 통한 정교한 모델 설계 지원</li> <li>▪ 정확도/재현율 및 기대수익 등 다양한 모델 평가 시각화 지원</li> </ul>	<ul style="list-style-type: none"> <li>▪ 리포팅을 통한 데이터 및 프로젝트 관련 다양한 모델링 비교 분석 기능</li> <li>▪ 최종 결과물을 유저가 원하는 형태로 export 지원</li> <li>▪ C/JAVA/Lib/수식 등 다양한 형태로 예측함수 추출 가능</li> </ul>

# 프로젝트 수행 및 적용 Process

**고도화 예측모형 생성 Project(알고리즘 최적화 및 시스템 개발 기간 1개월)** : 솔리드웨어의 ML Scientist가 DAVinCI Core를 활용, 고객사의 프로젝트 목적에 맞는 Customized Prediction Model(커스텀 예측모형)을 개발하여 고객사 데이터 환경과의 통합 수행



## 고객사 DB (BIG DATA)

금융, 소비재, 소매, 무역 등 다양한 분야의 과거 데이터 활용



DAVinCI LABS

## DAVinCI Core

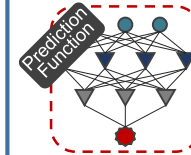
(DAVinCI LABS 분석 엔진)



## Data Scientist

(Solidware)

DAVinCI Core가 탑재된 솔리드웨어 서버를 활용, 솔리드웨어 데이터 사이언티스트의 고객사 DB 분석 수행



고객사의 프로젝트 목적에 최적화 된 예측모형 생성



머신러닝 알고리즘 기반 커스터마이징 된 전처리 기능 등 고객사 환경에 특화된 맞춤형 솔루션 구축

**Project 종료 이후 운용 (DAVinCI LABS Integration)** : 머신러닝 사이언티스트의 노하우가 담긴 Customized DAVinCI LABS를 통해 기존의 신용평가 모델 사용과 더불어 다양한 예측 분석에 활용 가능



## 신규 데이터 유입 (NEW DATA)

신규 데이터 발생/투입 시 DAVinCI LABS가 데이터의 변동을 자동적으로 반영하여 예측모형 Update 실시



DAVinCI LABS

## DAVinCI Studio

(DAVinCI LABS 시각화 엔진)



## 현업 실무자

(Client)

솔리드웨어 머신러닝 사이언티스트의 노하우 및 사고 흐름을 DAVinCI LABS를 통해 고객사 현업 부서에 내재화, 분석 솔루션의 범용적 활용 기반 마련



타겟 마케팅



조기경보



투자운용

DAVinCI LABS를 활용, 지속 가능한 예측분석 가능

고객사의 환경 및 니즈에 맞춰 통합 된 DAVinCI LABS를 활용, 신용평가 외의 예측(타겟 마케팅, 이탈 감지 등) 모형을 자체적으로 생성 가능

# 유지보수 방안 및 지원 체계

예측 모형 업데이트 작업 필요 시 자체적으로 DAVinCI LABS를 통해 유지보수 작업 가능

✓ 모형 업데이트 작업 필요 요건사항:

1. 신규 변수 추가
2. 기존 변수 특성 퇴색
3. 데이터 양 추가
4. 고객군 Trend 변화
5. 고객사 내부 정책 변경
6. DAVinCI LABS 내부 기능(Module) 추가



## DAVinCI LABS를 통한 유지보수 관련 장점

1. 자체적인 유지보수를 통해 수행 인력 관련 비용 대폭 절감 (필요 시 On-Demand 서포트 제공 가능)
2. 고객사와 제안사간의 불필요한 커뮤니케이션을 배제하여 업데이트 작업 일정 최소화
3. 즉각적인 모형 업데이트를 통한 사업 프로세스 지체 요소 제거

# 상세 기능별 Use Case Example

## ◆ Case I



**Credit Scoring**

### 신용평가 (Credit Scoring)

개요	신규 금융상품 셀링을 위한 타겟 고객 신용도 평가 지표 산출
업종	각종 금융사
부서	심사팀/리스크관리팀
목적	산출 지표를 통한 채무 불이행 고객 예측 및 대응

## ◆ Case II



**Target Marketing**

### 타겟 마케팅 (Target Marketing)

개요	고객의 라이프 스타일/스테이지 등 패턴 분석을 통한 유효 타겟 집단 산출
업종	각종 서비스 업
부서	마케팅/홍보팀
목적	타겟 고객집단 특성 도출을 통한 소구 포인트 파악과 매출 drive 전략 수립

## ◆ Case II



**Churn Detection**

### 고객 이탈 예측 (Churn Detection)

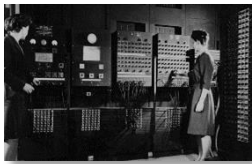
개요	고객 retention을 위한 핵심 지표 산출
업종	각종 서비스 업
부서	영업
목적	산출 지표를 통한 이상 고객 집중 관리 전략 수립

# Vision

“애플이 초기 컴퓨터 기술과 일반 사용자를 이어주는 다리가 되었다면,  
솔리드웨어는 인공지능 기반 분석 기술과 일반 사용자를 연결하는 것을 지향”

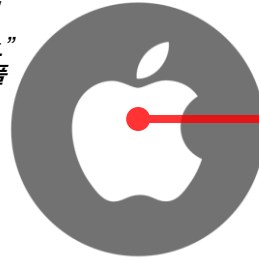
## 초기 컴퓨터 기술

- 목적: 과학/군사용
- 개발: 연구원/전문가 집단



## Apple Eco-System

“미국의 2억 3천 5백만 인구 중,  
오직 극소수의 사람만이  
컴퓨터를 사용할 수 있다.”  
- 1983, 애플



“Think different.”



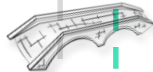
## 사용자

- 대상: 일반 사용자(개인/기업/대학)



## 인공지능 기술

- 목적: 과학/기타
- 개발: 학계/연구소/거대IT기업



“Your Personal Artificial Intelligence”

데이터 전처리

분석  
시각화  
리포트



리스크 관리

타겟 마케팅

의사결정 최적화

Etc.

운영 관리 자동화

OCR

Neural Network

Etc.

GBM

VTT

ETL

Regression Forest

## 사용자

- 대상: 일반 사용자(개인/기업(중소규모포함)/공공기관/대학)







# Solidware

We turn data into money.

# Thank you

For more information, please contact at

[sophie.eom@solidware.io](mailto:sophie.eom@solidware.io)

or 010 4008 5521 (엄수원, 공동대표이사)