

SPM[®] 8.2

Salford Predictive Modeler[®]

SPM[®] 8.2

SPM Salford Predictive Modeler[®] 소프트웨어 제품군은 데이터베이스 및 데이터 세트의 크기, 복잡성, 구조에 상관없이 예측, 설명, 분석 모델을 개발할 수 있는 매우 정확한, 초고속의 플랫폼입니다.

SPM 소프트웨어 제품군의 데이터 마이닝 기술에는 분류 분석, 회귀 분석, 생존 분석, 결측값 분석, 히스토그램 및 클러스터링/세분화가 포함되어 있어 데이터 과학 프로젝트의 모든 측면을 다룰 수 있습니다. SPM 알고리즘은 정교한 데이터 과학 분야에서 필수적입니다.

기능

- 선도적인 모델 분석자의 작업 방식을 패키지로 구현한 자동화 시나리오(70개 이상)
- 분석자가 모델 개발의 창의적인 측면에 집중할 수 있도록 단순 반복 작업 최소화
- 대규모 데이터 세트에 적용할 수 있는 회귀분석, 분류분석, 로지스틱 회귀분석



SPM 소프트웨어 제품군의 자동화는 모델 탐색 및 정제 프로세스의 상당 부분을 수행함으로써 분석자가 모델을 빠르게 구축할 수 있도록 합니다. 분석자가 항상 모든 권한을 가지고 있지만, SPM 소프트웨어 제품군의 자동화는 분석자의 다음 단계를 예상하고 분석자가 쉽게 검토할 수 있도록 대체 모델링 전략의 결과를 패키지로 제공합니다.

SPM® 8.2 Product Versions

ULTRA: The best of the best. 앙상블 모델링, 교호 효과 탐지 및 자동화의 주요 고급 기능이 있습니다. 최첨단 기술과 빠른 실행 시간이 필요한 모델 작성자에게 적합합니다.

PROEX: 최첨단 데이터 마이닝 기술이 필요한 모델 작성자에게 적합합니다. 숙련된 데이터 분석가를 위한 전형적인 모델링 실험의 자동화 기능과 Salford 데이터 마이닝 엔진의 수십 가지 확장 기능이 있습니다.

PRO: 예측 모델링 워크벤치로서 전문 데이터 마이너에게 적합합니다. 기존 통계 모델링 도구, 프로그래밍 언어, 리포팅 서비스 및 워크플로우 자동화 옵션 선택을 지원하는 다양한 기능이 있습니다.

BASIC: 기본 제품. Salford Systems는 수상 경력에 빛나는 데이터 마이닝 엔진입니다. 확장 또는 자동화 또는 주변 통계 서비스, 프로그래밍 언어 및 정교한 보고 기능은 없습니다. 적은 예산으로 세계적으로 유명한 엔진을 사용할 수 있습니다.

“SPM 소프트웨어 제품군을 이용하여, 1주일의 작업 소요 시간을 1일로 단축하세요!”

Features

SPM® 8.2

- > 성능 향상
- > 기술 통계량 관리
- > 모델 기반의 대체 또는 요약 통계량을 사용한 대체
- > 상관 분석 및 다차원 스케일링
- > 정확한 테스트 파티셔닝 부분
- > KEEP 및 EXCLUDE 제어
- > 성능 통계량의 확장: ROC 분산 및 Kolmogorov-Smirnov (K-5) 통계량
- > 새로운 시계열 지원 메커니즘
- > 새로운 자동화된 모델 실험
- > 새로운 점수 분포 제어 및 시각화
- > 재귀 기능 제거

CART®

Classification and Regression Trees

CART® 소프트웨어는 고급 분석 분야에 혁명을 일으키고 데이터 과학의 현 시대를 열어 준 최고의 분류 트리입니다. CART는 최신 데이터 마이닝에서 가장 중요한 도구 중 하나입니다.

CART는 정확성, 성능, 기능 세트, 기본 제공 자동화 및 사용 용이성 면에서 독보적인 존재입니다.

CART는 전문가와 비전문가 모두가 사용할 수 있도록 만들어졌으며, 다른 분석 도구를 사용한다면 숨겨져 있을 수 있는, 데이터 간의 중요한 관계들을 신속하게 나타낼 수 있습니다.

기술적으로, CART는 스탠포드 대학과 버클리 캘리포니아 대학에서 세계적으로 유명한 4 명의 통계학자들이 1984년에 소개한 획기적인 수학 이론을 기반으로 합니다. Salford Systems의 CART는 원 저작 코드를 구현하는 유일한 의사 결정 트리 소프트웨어입니다. CART 이론을 처음 소개한 사람들은 계속 Salford Systems와 협력하여 CART를 지속적으로 향상시킵니다.

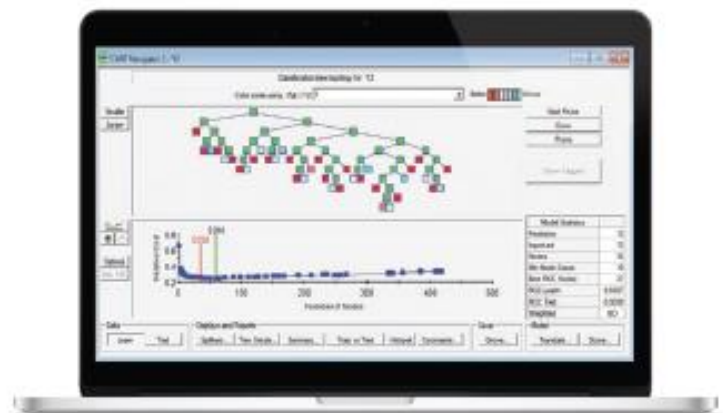
CART는 특히 시장 조사 및 웹 분석 결과를 향상시키도록 확장되었습니다.

CART는 고속 전개를 지원하여, Salford Systems이 대용량 데이터를 실시간으로 예측하고 점수 매길 수 있습니다.

Newest Features

SPM® 8.2

- 이상징후 감지를 위한 자율 학습
- 핫스팟을 사용한 연관성 분석
- 강제 분할
- 하위 노드의 확장 가능성 제한
- 자동화된 사전 검색
- 자동화된 바이어스 패널티



TreeNet®

Gradient Boosting

TreeNet® 소프트웨어는 Salford Systems의 가장 유연하고 강력한 데이터 마이닝 도구로, 2002년에 출시된 이후 주요 데이터 마이닝 대회에서 최소 12개의 상을 수상했습니다.

최고의 정확도

TreeNet은 단일 모델이나 배경 또는 기존 부스팅과 같은 앙상블로는 일반적으로 얻을 수 없는 정확도가 장점입니다. 신경망과는 달리, TreeNet은 데이터 오류에 민감하지 않으므로 시간이 많이 소요되는 데이터 준비, 사전 처리 또는 결측치의 대체가 필요하지 않습니다.

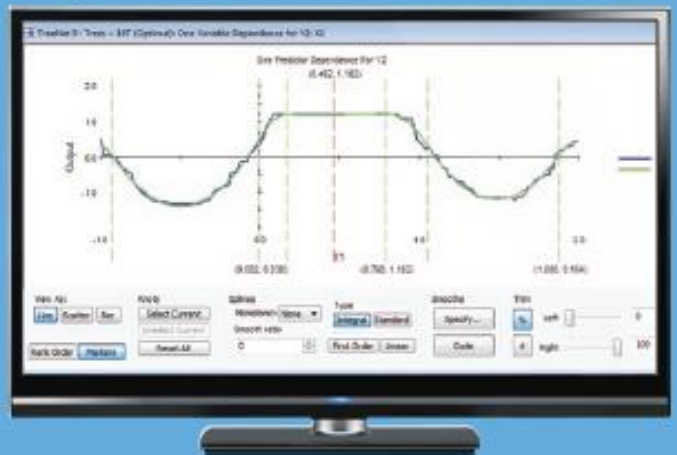
TreeNet은 목표 레이블이 잘못된 데이터에도 견고합니다. 이러한 유형의 데이터 오류는 기존의 데이터 마이닝 방식으로는 처리되기 매우 어려우며 기존의 부스팅에는 치명적일 수 있습니다. 하지만, TreeNet은 기존 모델과 너무 차이가 큰 훈련(training) 데이터 포인트를 동적으로 거부하므로 이러한 오류에 일반적으로 영향받지 않습니다.

고급 기능

교호 효과 탐지 통계량은 예측 모델에 어떤 종류의 교호 작용이 필요한지 여부를 확인하고, 특히 어떤 교호 작용이 중요한지 알아내기 위해 검색합니다. 교호 작용 탐지 시스템은 모델 성능을 (때로는 극적으로) 향상시킬 뿐만 아니라 가치 있는 새로운 세그먼트와 이전에 인식하지 못했던 패턴을 발견하는데 도움이 됩니다.

NextGen SPM® 8.2

- 정기적인 Gradient Boosting™ (RBOOST)
- TreeNet으로 랜덤포레스트 구축
- RuleLearner® : TreeNet의 정확성과 회귀 분석 및 CART의 해석 가능성 및 투명성의 장점
- 새로운 손실 함수 differential lift
- 교호 효과 크기 리포팅: 가장 중요한 교호 작용을 발견, 측정하여 그래프 작성
- 고급 컨트롤 : subsampling and influence trimming
- Monotone constraints
- 다양한 예측 변수의 수
- 시작 값 지정
- TreeNet splines을 사용한 모델링



MARS®

Automatic Non-linear Regression

MARS® 소프트웨어는 비선형성 주효과 및 교호 효과를 찾아내면서 전통적인 회귀 분석과 유사한 형태의 결과를 선호하는 사용자에게 이상적입니다. 회귀 모델링에 대한 MARS의 접근 방식은 다른 회귀 분석 방법으로는 알아내기 어려운 중요한 데이터 패턴과 관계를 효과적으로 알 수 있습니다.

회귀 및 분류

MARS 모델은 통신사 고객의 월별 평균 청구액 또는 웹사이트 1회 방문 시 구매 예상액과 같은 수치를 예상합니다. MARS는 예 / 아니오 결과에 대한 고품질 분류 모델을 생성 할 수도 있습니다. MARS는 변수 선택, 변수 변환, 교호 작용 탐지 및 자체 테스트를 모두 자동으로 고속으로 수행합니다.

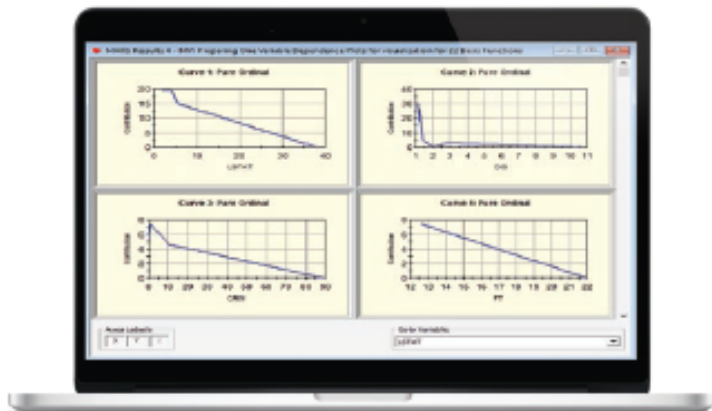
뛰어난 성능

MARS가 뛰어난 성능을 발휘하는 분야로는 발전 업체의 전기 수요 예측, 제품의 엔지니어링 사양에 대한 고객 만족도, 지리 정보 시스템 (GIS)의 유무 모델링 등이 있습니다.

Features

SPM® 8.2

- ▶ 기본 함수의 최대 수를 변화시키면서 모델 구축 (BASIS 자동화)
- ▶ 평활도를 변화시키면서 모델 구축 (MINSPAN 자동화).
- ▶ 교호 작용 차수를 변화시키면서 모델 구축 (INTERACTIONS 자동화).
- ▶ 모델링 속도를 변화시키면서 모델 구축 (Automate SPEED)
- ▶ 변수 추가에 대한 패널티 정도를 변화시키면서 모델 구축 (PENALTY MARS 자동화)



Random Forest®

Breiman and Cutler's Random Forests

Random Forests® 소프트웨어는 다중 대안 분석, 무작위 추출 전략 및 앙상블 학습을 활용하는 배깅 (bagging) 도구입니다. 그 강점은 데이터의 이상치 및 이상 징후를 발견하고, 군집을 표시하고, 미래의 결과를 예측하고, 중요한 예측변수를 식별하고, 결측치를 대체하고, 통찰력있는 그래픽을 제공하는 것입니다.

클러스터와 세그먼트

Random Forests는 나무가 자란 후에 적용되는 방법에 의해 대부분의 인사이트를 제공합니다. 변수의 중요성을 평가하는 새로운 방법뿐만 아니라 데이터의 클러스터 또는 세그먼트를 식별하는 새로운 기술이 있습니다. 이 방법은 University of California, Berkeley의 Leo Breiman과 Adele Cutler가 개발했으며 Salford Systems에게 독점적으로 라이선스되었습니다. Salford Systems는 Random Forests의 생존 공동 저자인 Adele Cutler 교수와 공동 연구를 진행하고 있습니다.

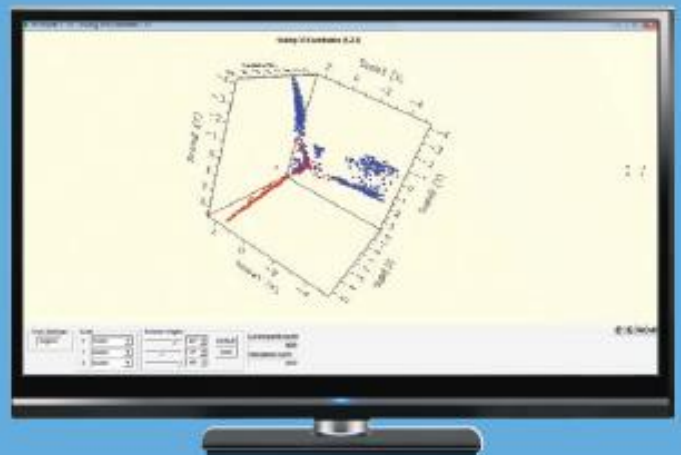
광범위한 데이터 세트에 적합

Random Forests는 서로 영향을 받지 않으며 구성된 여러 개의 CART 트리 모음입니다. 의사 결정 나무에서 만들어진 예측의 합이 전체 숲의 예측을 결정합니다. 랜덤 포레스트는 행의 수는 10,000개 미만이고 열의 수는 수백만 개인 중소 규모 데이터 세트의 복잡한 데이터 구조를 분석하는 데 가장 적합합니다.

Advantages

SPM® 8.2

- 자동화를 통한 파라미터 최적화
- 특수 RF 클러스터링 향상
- Extreme Trees: 임의 분할을 위한 특수 컨트롤
- 고급 결측치 대체 옵션
- Penalty Configuration: 특정 변수의 포함에 대한 패널티 부여 가능





회사 안내

Salford Systems는 예측 모델 개발의 모든 측면에서 데이터 사이언티스트를 지원하는 최첨단 기계 학습 기술 전문 기업입니다.

PEDIGREE

Salford Systems의 기술은 현대 기계 학습 및 데이터 마이닝의 원 창시자와 직접 협력하여 개발되고 향상되고 있습니다. 우리는 선도적인 대학들과의 협력하여 R & D 프로그램을 활발하게 운영하고 있습니다.

THE SOFTWARE

Salford의 도구는 사용 용이성, 대용량 데이터 작업 기능, 고속 모델 개발, 견고성 및 신뢰성, 매우 정확한 모델의 일관된 제공으로 잘 알려져 있습니다. Salford의 모델링 자동화 도구는 초보 데이터 사이언티스트에게 모델 개발의 복잡한 과정을 안내하고, 전문적인 데이터 사이언티스트들이 세계 정상 수준의 예측 모델을 개발하는 데 도움을 줍니다.

Salford Systems은 Salford Systems의 소프트웨어 및 알고리즘을 사용하는 고객과 함께 산업계에서 인정 받고 있습니다.

CUSTOMER AWARDS AND ACHIEVEMENTS

PAKDD 데이터 마이닝 학술대회, the annual Pacific-Asia Knowledge Discovery and Data Mining data mining competition

- 교차 판매 과제, 재무 데이터 세트

DMA Direct Marketing Association Analytics Challenge

- 맞춤형 고객 서비스를 제공하기 위해 고객 총 가치 예측
- 의료 대응 과제
- 후원을 중지한 기부자 세그먼트에 대한 Make-A-Wish 재단의 맞춤 솔루션
- 표적 마케팅 과제

SALFORD AWARDS AND ACHIEVEMENTS

지식발견 및 데이터마이닝(KDD : Knowledge Discovery and Data Mining) 국제학술대회

- 웹 분석 업무

SPM의 알고리즘 작성자인 Drs. Leo Breiman 와 Jerome Friedman의 SIG KDD Innovation Award 수상

- 상용 시스템의 이론 및/또는 개발 발전에 지속적인 영향을 미친 데이터 및 데이터 마이닝 관련 지식 발견 분야에 큰 기술적 기여

Duke 대학의 고객관계관리를 위한 Teradata 센터

- 고객 이탈 모델링, CRM

CONTACT

경기도 안양시 동안구 시민대로 401 (관양동, 대륭테크노타운 15차 901호)

전화: 031-345-1170 | 팩스: 031-345-1199

<http://www.datalabs.co.kr/>

