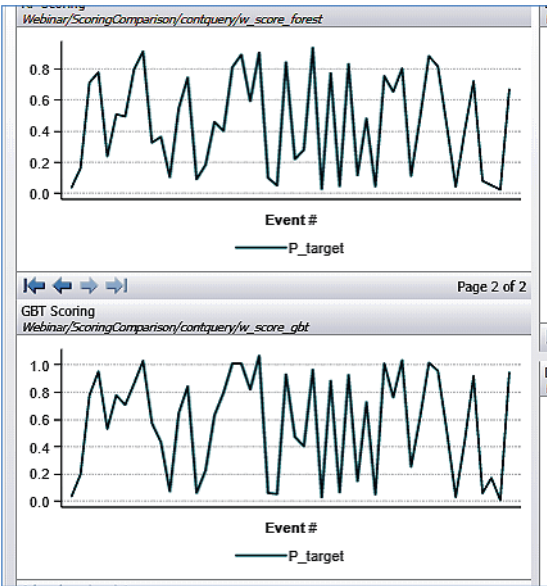


SAS® Event Stream Processing

지속적으로 발생하는 빅데이터에 대한 실시간 분석



SAS® Event Stream Processing은 어떤 솔루션인가?

SAS Event Stream Processing은 초당 수백만건 이상 발생하는 이벤트를 감지하고 이의 패턴을 분석하여 올바른 대응 조치와 발행할 경보, 저장할 데이터, 중요하지 않은 이벤트 등을 구분합니다.

SAS® Event Stream Processing이 중요한 이유는?

데이터가 전송 중일 때도 센서, 장치, 전송, 운영 트랜잭션 등에서 발생하는 스트리밍 데이터를 통해 가치를 창출합니다. 또한 기본 제공되는 연산자, 함수, 루틴 및 고급 분석 기능을 사용하여 1/1000초 이내의 빠른 응답 속도로 데이터 스트림을 측정합니다. 이에 따라 사용자는 이벤트에 대한 인사이트를 획득하여 적절한 대응을 할 수 있습니다.

SAS® Event Stream Processing는 누구를 위한 솔루션인가?

설정 가능하고 내제된 엔진을 탑재한 소프트웨어로 응용 프로그램 개발자, IT 엔지니어, 시스템 아키텍처를 위한 제품입니다. 웹 기반 인터페이스로 데이터 조작 및 프로젝트 구축이 용이하며 대시보드에 표시되는 시각화된 정보를 통해 결과를 손쉽게 테스트 및 검증할 수 있습니다.

빅데이터 세계에서 경쟁력을 유지하려면 보다 민첩하게 업무를 처리하여 신속한 의사 결정을 내릴 수 있어야 합니다. 특히, 사물인터넷 데이터의 급속한 성장에 대처하기 위해서는 지속적으로 변화하는 상황을 파악하여 관심 이벤트가 발생할 때 상황에 맞는 적절한 조치를 취해야 합니다. 이벤트 스트림 프로세싱 기술을 사용하면 데이터 수집 및 처리에 영향을 주지 않습니다. 지속적인 쿼리가 항상 라이브 데이터를 분석하여 관심 패턴을 탐지하기 때문에 사용자가 실시간으로 응답할 수 있습니다.

SAS® Event Stream Processing는 스트림 이벤트를 파악할 수 있게 해주는 솔루션으로, 전통적인 '스트림, 스코어, 스토어(stream, score and store)' 모델과 달리 쿼리를 저장하고 이를 통해 데이터를 스트림 처리합니다. 이 경우 데이터가 수신되는 즉시 분석되며, 새로운 이벤트가 발생하면 상황에 따라 인텔리전스를 업데이트할 수 있습니다. 특히 1/1000초 미만의 높은 처리속도의 유연성이 뛰어난 프로세싱 모델에서 최적의 성능을 발휘하도록 설계되었습니다.

SAS® Event Stream Processing 특징

- **실시간 정보 제공 및 결과 산출** 실시간 처리에는 운영, 트랜잭션, 시스템 IoT 센서 및 장치 등에서 나오는 스트리밍 데이터가 중요합니다. 데이터나 형식에 관계없이 SAS는 고객 경험, 자산 및 IT 성능, 사기, 위험, 보안 등의 정보 기반 애플리케이션에서 스트리밍 데이터의 가치를 즉시 높여줍니다.
- **강력한 실시간 데이터 스트림 프로세싱** 강력한 스트림 프로세싱은 데이터가 저장되기 전에 이를 처리하여 빅데이터 프로세싱 비용을 절감합니다. 연속 쿼리가 유연한 스레드 프로세싱 모델에서 보다 신속하게 실행되어 데이터가 스트림 처리되는 동안 기본적으로 동적 업데이트, 삭제 및 삽입을 지원합니다.
- **적합한 데이터 저장** 더욱 빠르고 효과적이며 강력한 스트리밍 데이터 분석을 통해 데이터 변환을 데이터 저장 이전에 실행 할 수 있습니다. 이를 통해 데이터 저장 및 변환 비용을 절감할 수 있습니다. 유연한 스레드 프로세싱 모델에서 빠른 속도로 연속 쿼리를 실행하여 데이터가 전송 중일 때도 중요한 비즈니스 데이터를 지능적이고 능동적인 방식으로 필터링, 감지, 업데이트, 삭제 및 추가합니다.
- **정확한 분석을 통한 최적의 의사 결정** 정확한 데이터는 올바른 의사결정의 기반이 됩니다. SAS® Event Stream Processing은 인스트림의 기본인 데이터 품질 연산자, 자연 언어 텍스트 처리 함수 및 풍부한 고급 분석 알고리즘을 제공하여 가장 복잡한 이벤트 스트림 패턴까지 탐지해 냅니다.
- **정교한 이벤트 스트림 분석으로 변화에 신속하게 대응** 데이터 보유량과 이동속도 및 연결 데이터 출처와 관계없이 직관적인 단일 인터페이스를 통해 모든 분석 과정을 간소화하여 쉽게 관리할 수 있습니다. 또한 SAS에서 정의한 통합 이벤트 스트림 데이터 명령과 기타 프로그래밍 언어로 빠르게 변화하는 데이터에 보다 적절히 대응할 수 있습니다.

SAS® Event Stream Processing 주요 기능

SAS Event Stream Processing은 대용량의 스트리밍 데이터를 빠르게 수집하여 초당 수백만 건의 이벤트를 분석하므로 사용자는 전송 중인 데이터의 이벤트를 파악할 수 있습니다. 빠른 속도로 이동하는 대용량의 데이터를 원활하게 처리하여 사용자는 종단 장치(edge device), 데이터 센터, 클라우드 등 전체 환경에서 수집되는 IoT 데이터를 통합, 시각화, 변환 및 분석할 수 있습니다. 단, 하드웨어 환경에 따라 처리속도가 달라질 수 있습니다.

게시/구독 아키텍처를 구성하는 어댑터와 커넥터를 통해 수신 데이터를 판독합니다. 이렇게 수신되는 이벤트 데이터는 이벤트 스트림 프로세서 창에 게시됩니다. 시각적 인터페이스에서 창(windows), 프로시저 및 연산

자를 쉽게 정의할 수 있기 때문에 데이터를 스트리밍하는 연속 쿼리 역시 간단하게 정의할 수 있습니다. 스트리밍 데이터의 패턴을 검사하고 지능적 필터링을 통해 이상점을 발견하면, 이를 저장하여 심층적인 조사를 하거나 연관성이 없는 데이터는 삭제합니다. 다운스트림 응용 프로그램은 사전에 설정된 작업으로 스트리밍 분석 결과를 수신합니다. 이러한 방식을 통해 변화하는 환경에 빠르게 대응하여 새로운 IoT 기회를 포착할 수 있습니다.

센서 또는 클라우드 환경의 스트리밍 데이터 사용 및 연결
대부분의 데이터 센터에서는 빠르게 증가하는 IoT 데이터를 처리할 준비가 되어있다거나 엣지(edge)와 포그(fog) 및 클라우드 데이터 수집에 대한 준비가 되어있지 않습니다. 현실은 웹 트랜잭션, 센서 장치 및 운영 시

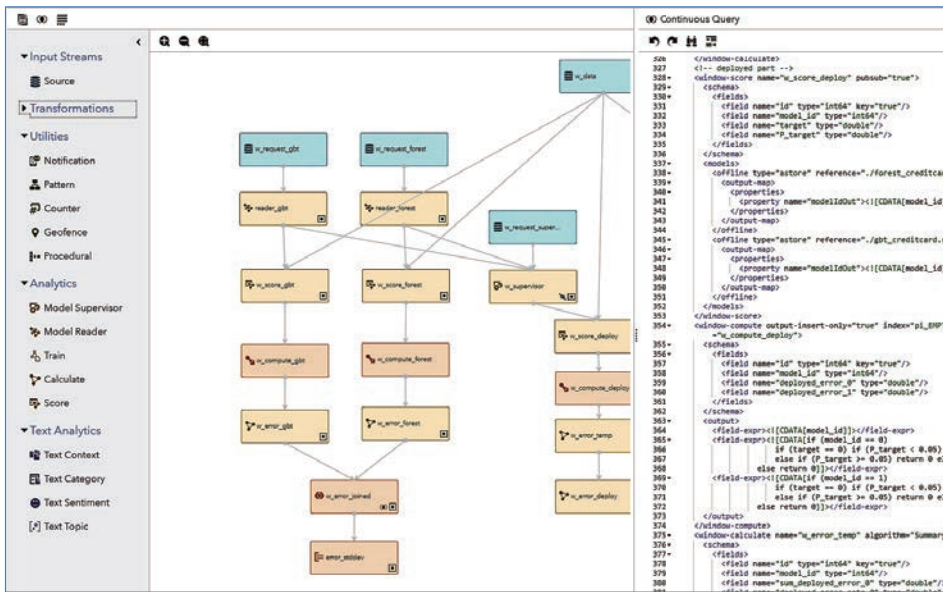


그림 1: SAS Event Stream Processing의 설계 인터페이스를 통해 모든 고급 분석 및 스트림 프로세싱 창에 액세스할 수 있습니다.

Name	Type	Tags	Modified Date	Modified Time	Modified By
PMU_Transpose	ESP Project		2017-11-30	11:46	anonymous
Sample_Cluster_Mod	ESP Project		2017-11-29	13:24	anonymous
Truck	ESP Project	TruckFleet, Monitoring, IOT	2017-11-28	15:15	anonymous
pg_usecase2	ESP Project		2017-11-27	15:04	anonymous
project	ESP Project	kmeans, demo	2017-11-15	12:37	anonymous
project	ESP Project	Demo, DBSCAN	2017-11-15	12:34	anonymous
Sample_PMU	ESP Project		2017-11-09	11:33	anonymous
ScoringComparison	ESP Project	RandomForest, GradientBoost, ML	2017-11-01	10:29	anonymous
Clustering_Model	ESP Project		2017-11-01	10:28	anonymous
annotationProject	ESP Project	image	2017-10-18	18:22	anonymous
Sample_Cluster	ESP Project		2017-10-17	15:52	anonymous
project	ESP Project	kmeans	2017-09-28	09:46	anonymous
project	ESP Project	dbscan	2017-09-28	09:46	anonymous
HealthMonitoring	ESP Project	health, wearables	2017-09-28	09:44	anonymous
SocialMedia	ESP Project	sentiment, social	2017-09-28	09:33	anonymous

Name	Tags	Modified Date	Modified Time	Modified...
Solace03		2017-11-01	10:47	anonymous

Identification
ID: d5e39d8b-cd39-4608-9bc4-80149b21d4ef

그림 2: SAS Event Stream Processing 스튜디오는 간편한 프로젝트 리포지토리 탐색 기능을 제공합니다.

시스템에서 발생하는 데이터 스트림조차도 통합 및 활용이 쉽지 않습니다.

SAS® Event Stream Processing은 데이터 스트림의 다양한 형식을 사용할 수 있는 강력한 기능을 갖추고 있어 분석 결과를 기존 시스템이나 애플리케이션에 쉽게 통합할 수 있습니다. 기본 제공되는 어댑터와 커넥터의 광범위한 제품군을 통해 사용자는 라이브 데이터 스트림을 수집하고 조작하여 관심 패턴의 탐지, 필터링, 집계, 연관성 제공, 설명 등을 할 수 있습니다. 이는 정형 및 비정형 텍스트 데이터에도 적용되며, 결과가 다른 시스템에 전송되어 효율성을 개선하고 비용을 절감합니다. 기업은 현재 보유하고 있는 기술을 향후에도 사용할 수 있다는 확신을 가지고 현재 필요로 하는 데이터 소스를 사용할 수 있습니다.

- 정형/비정형 데이터의 실시간 데이터 스트림을 게시 및 구독할 수 있는 광범위한 데이터 어댑터 및 커넥터
- 게시 및 구독이 기본 제공되는 어댑터: Hadoop (HDFS Yarn), Apache Camel, OS/soft PI, Axeda, Rabbit MQ, Solace, Tervela Data Fabric, XML/JSON File Socket Adapter, SAS® LASR™ Analytic Server, IBM DB2, IBM Netezza, IBM WebSphere MQ, SAP Sybase ASE, Tibco Rendezvous, JMS, File/Socket, Database ODBC 및 SAS 데이터 세트
- 사용자 정의 가능한 게시/구독 API는 C 또는 Java로 작성 가능
- 게시만 가능: Twitter, Log Sniffers (Oracle, Greenplum), Network Sniffer, Google Protobuf Sniffer, SYSLOG, HTTP RESTful 인터페이스 (SOAP 및 SMTP에는 구독만 가능)
- 이벤트 핸들러를 사용하여 복수의 입력 스트림을 관리
- 보장 모드의 기능 또는 게시/구독 API를 통해 이벤트 스트림을 지속적으로 동기화하여 이벤트 스트림의 지속성 확인

고급 스트리밍 분석, 유연한 시각적 개발 환경

SAS® Event Stream Processing은 이벤트 연관성을 파악할 수 있는 SAS 고급 분석 알고리즘과 규칙을 사용하는 모듈식의 연속 쿼리를 구축하기 위해 시각화 기능이 뛰어난 인터랙티브 인터페이스를 제공합니다. 시스템 아키텍처와 애플리케이션 개발자는 창과 커넥터의 팔레트에 포인트 앤클릭 (point-and-click) 액세스가 가능하여 복잡한 스트리밍 분석 모델을 쉽게 설계할 수 있습니다. 또한 인터랙티브 테스트 모드를 사용하여 결과 배포 전에 논리를 평가하고 결과를 검사할 수 있습니다. 이벤트는 프로젝트 전반에 걸쳐 공유될 수 있고, 이전의 활동이 저장되어 시각적인 드래그 앤드 드롭 (drag-and-drop) 환경에서 현재 이벤트와 비교 가능합니다. 즉, 움직이는 데이터를 쉽게 정의, 업데이트 및 수정하여 빠르게 등장하는 새 이벤트를 처리할 수 있습니다.

실시간 고급 분석은 높은 빈도로 사용되는 계산과 고급 클러스터링 알고리즘을 제공합니다. 고급 클러스터링 알고리즘은 저장 데이터를 통해 트레이닝되고 스트리밍 데이터를 통해 배포됩니다. 알고리즘을 지원하는 분석 기법으로는 SVDD, 로버스트 주성분 분석(RPCA), 랜덤 포레스트, 그래디언트 부스팅 및 스트리밍 회귀 분석 등이 있습니다.

- 포인트앤클릭 설계와 이벤트 로직 테스트를 위한 풍부한 그래픽 인터페이스
- 이벤트 모델 정의를 위한 비교 가능한 C++ 라이브러리와 XML 구문 사용 가능
- 유연한 모듈식의 창 기반 아키텍처로 복잡한 연속 쿼리 정의
- 패턴, 상관 관계, 계산 및 집계의 무제한적 수 탐지
- 절차적, 선언적 등의 연산자, 규칙 및 사용자 정의 프로시저가 C++, DS2 또는 SAS DATA 스텝 콜아웃으로 정의
- 연속 쿼리는 연결된 창 유형의 시리즈로 정의
- 기본 제공되는 일반 데이터 품질 루틴은 저장되기 전에 라이브 스트림 데이터를 정리, 표준화 및 필터링하는 데 사용하여 다운스트림 처리 감소
- 인스트림 머신 러닝 K-평균 클러스터와 라이브 스트림 분석 채점으로 거의 무한대의 고급 분석 계산 포함
- 삽입, 삭제, 교체 등 활성 스트림 처리에 대한 동적 업데이트로 이벤트 스트림 평가 지원
- 게이트웨이, 장치, 기존의 모든 C++ 애플리케이션에 대한 뛰어난 임베디드 기능

고급 패턴 매칭 알고리즘

스트림 내에서의 데이터 프로세싱이 그 어느 때보다 철저해집니다. SAS® Event Stream Processing은 적용 즉시 사용 가능한 파싱, 필터, 조인, 필드 계산, 패턴 매칭 등의 함수를 제공합니다. 여기에는 변환, 정규화, 매칭, 확인 등 루틴 데이터 관리 작업을 위해 특별히 설계된 함수들이 포함되어 있습니다. 이 외에도 기본 제공되는 함수를 편집하고 커스텀 함수를 정의할 수 있어 유연성이 향상됩니다. 비정형 텍스트 처리 함수는 개념, 실체, 사실들의 자연어 처리 (NLP) 추출, 텍스트 분류와センチメント 식별 등을 지원합니다. 예측 알고리즘이 향후 이벤트를 예측하고 고급 관심 패턴을 생성하는 한편, 인스트림 클러스터링은 즉석에서 같은 종류의 이벤트 그룹을 만듭니다. 단일 쿼리에서 순차 및 시간 기반 이벤트 등 여러 이벤트를 검사할 수도 있습니다.

- 머신 러닝 스트리밍 알고리즘이 지원되면서 다양한 학습 알고리즘에 따라 스코어링 및 학습 프로시저 창을 생성하여 밀도 기반 군집(DBSCAN) 및 K-평균 군집을 추가 가능
- SAS Event Stream Processing과 함께 스트리밍 요약(단변량 통계), 피어슨 상관관계, 분포 적합성, 텍스트 토큰화, 세분화 상관관계, 그리고 단기 푸리에 변환이 패키지로 제공되는 인스트림 분석
- SAS Event Stream Processing과 함께 로버스트 주성분 분석(RPCA), 베이지안 네트워크, 컨볼루션 신경망 및 심층 신경망이 패키지로 제공되는 오프라인 트레이닝 알고리즘
- DBSCAN을 통해 군집을 자율적으로 구분. 트레이닝 창과 스코어 창을 함께 사용하여 DBSCAN 모델을 주기적으로 업데이트
- 지오펜스(geofence) 창을 통해 지오펜스 경계에 대한 객체의 위치 추적 가능. 객체가 지정된 지오펜스 경계로 접근하거나 이를 넘는 경우, 경보 메시지가 전송되고 이후 지오펜스 경계 내에서 이동 추적

- 정확한 모델 개발을 위해 과거 데이터에 기반하여 모델 트레이닝하고 이를 통해 정확한 데이터 분석
- SAS 모델 관리 기능이 통합되어 분석 모델과 SAS Event Stream Processing 프로젝트를 빠르게 자동 통합
- 유연한 모듈식 창 기반 아키텍처로 복잡한 연속 쿼리 정의
 - 호환 가능하고 광범위한 창 유형과 연산자를 바탕으로 패턴과 상관관계, 연산 및 집계 감지
 - 기본적으로 제공되는 공통 데이터 품질 관리를 사용하여 저장하기 전에 실시간 데이터를 정제, 표준화 및 필터링함으로써 다운스트림 프로세싱을 최소화
 - 분석하고자 하는 패턴에는 인스트림 기계 학습, K-평균 군집 및 라이브 스트림 분석 스코어링을 비롯하여 다양한 고급 분석 계산이 포함
 - 분석 모델에는 SAS DATA STEP, DS2, Python 및 써드 파티 코드 조각 포함
- 게이트웨이와 장치 및 기존의 C++ 응용 프로그램에 대한 뛰어난 내제 기능(전용 스트레드 풀 프로세싱 사용)

대시보드로 확인하는 라이브 이벤트 스트림

SAS Event Stream Processing의 스트림 뷰어는 라이브 이벤트 스트림을 한 곳에서 모니터링할 수 있는 실시간 대시보드로, 이를 통한 작업으로는 SAS 그래프 시각화 생성을 비롯해 각 SAS Event Stream Processing 서버에 있는 여러 개의 모델 보기, 대시보드 공유, 내제된 대시보드 생성이 있습니다. 또한 SAS Visual Analytics 리포트에 대한 링크를 실시간 대시보드에 삽입하여 이력 데이터에 대한 심층적 인사이트에 빠르게 접근할 수 있습니다.

- 데이터 스트림 지원:
 - Hortonworks DataFlow(HDF) NiFi 통합: SAS와 HDF가 즉각적인 스트리밍 및 심층 인텔리전스 제공
 - MapR 스트림 지원
 - IoT 커넥터 및 어댑터(OPC-UA, Twitter, Kafka, Flume, MQTT, 게이트웨이, 장치, Cassandra(어댑터에 전용)), BoardReader

- SAS Event Stream Processing for SAS® Viya® 및 SAS Cloud Analytic Services(CAS):
 - 스트리밍 분석을 위해 SAS Viya 데이터 관리/머신 러닝 모델을 SAS Event Stream Processing에 배포
 - SAS CAS에서 데이터가 전송 및 수신되는 동안 패턴을 스트리밍, 분석, 필터링 하여 파악하고 데이터 기록

확장 가능한 최적화된 프로세싱

이벤트 스트림 프로세싱을 사용하려면 시스템이 항상 온라인에 연결되어 있어야 하고 최적의 성능으로 매우 빠르게 처리할 수 있어야 합니다. 많은 양을 적은 시간 내에 처리하는 SAS® Event Stream Processing 소프트웨어는 범용 하드웨어를 사용하는 다른 스트림 프로세싱 엔진의 성능을 능가합니다. 또한 데이터량이 많아짐에 따라 분산된 인메모리 그리드 처리를 선형으로 확장하여 하드웨어 투자를 최적화합니다.

- SAS Event Stream Processing을 IoT 애플리케이션의 엣지 영역에 배포. 엣지 영역 계산용 SAS Event Stream Processing은 작은 용량의 중단 장치에 쉽게 배포할 수 있도록 사용 디스크 공간 축소
- SAS Event Stream Processing Python 개발 인터페이스:
 - 친숙하고 유연한 개방형 Python 인터페이스를 통해 개발 시간을 단축하고 이벤트를 개발, 게시 및 테스트한 후 SAS Event Stream Processing 프로젝트를 통해 스트리밍
- Python 게시/구독 API:
 - Python을 사용하여 이벤트를 게시하고 SAS Event Stream Processing 구독
- SAS Event Stream Processing의 언어를 사용(인 프로세스 이벤트 스트림 핸들러):
 - DS2, Python 및 C/C++
- BOSH Cloud Foundry: 대규모 분산 서비스에 사용

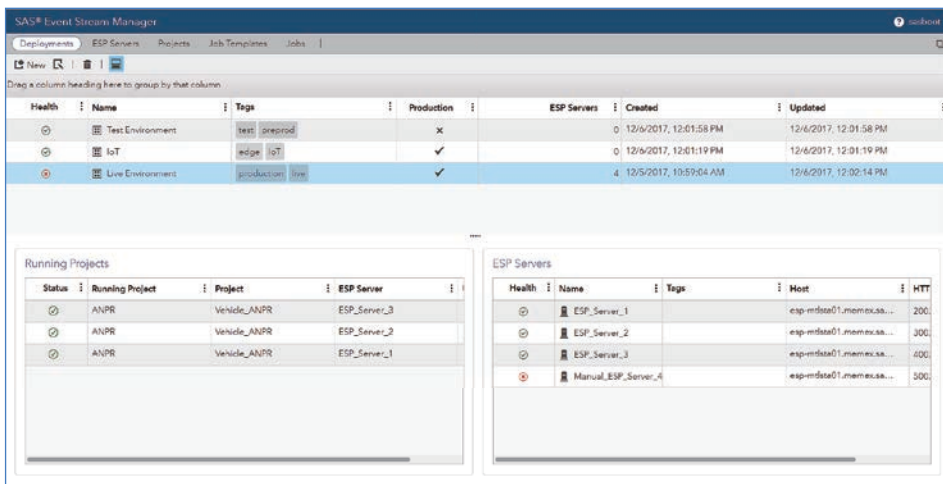


그림 3: SAS Event Stream Manager는 운영 스트리밍 환경을 위한 배포 시간을 줄여주며 IT 및 데이터 사이언티스트 사이의 협업을 지원합니다.

엔터프라이즈 관리

프로젝트 관리, 프로세싱 최적화 및 유지보수에 대한 IT 부서의 통제력을 강화하려면 이벤트 스트림 프로세싱 작동 상태를 중앙에서 관리하는 것이 매우 중요합니다. SAS® Event Stream Processing은 비즈니스 요구 기준에 부합하는 투명한 거버넌스를 제공하고 SAS 및 기타 소스에서 스트림 프로세서와 알고리즘을 통합할 수 있습니다. 데이터 크기가 가용 메모리 한계를 초과하는 경우에는 캐싱 스토어 (caching store) 를 사용할 수도 있습니다.

스트리밍 분석으로 다른 SAS® 솔루션과 통합

SAS® Event Stream Processing는 SAS® Customer Intelligence, SAS® Asset Performance Analytics, SAS® Security Intelligence 등 다른 SAS 솔루션의 결과를 사용합니다. 이러한 영역별 애플리케이션은 관심 패턴을 식별하여 SAS® Event Stream Processing에 임베드하는 데 사용 가능합니다. 스트림에 이상점이 있으면 정확한 식별을 위해 SAS 솔루션에 다시 보낼 수 있어, 모니터링할 새로운 관심 패턴을 정의하게 됩니다. 다단계 분석은 통합 분석 주기에 대한 끊임없는 개선을 통해 지속적인 모니터링을 보장합니다.

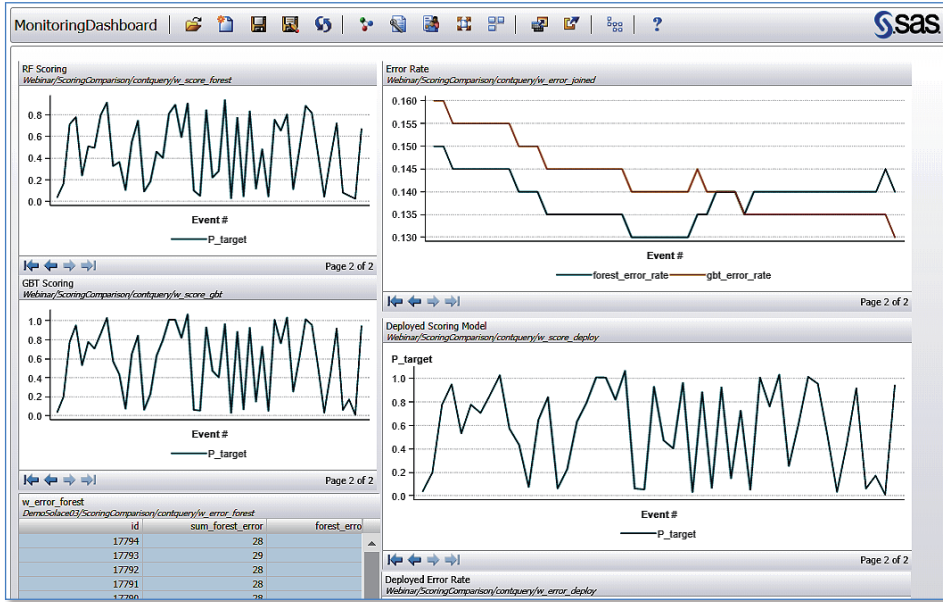


그림 4: 스트림 뷰어는 스트리밍 데이터 인사이트의 생성, 편집 및 공유용 실시간 대시보드입니다.

더 자세한 내용은 sas.com/korea/esp에서 확인하실 수 있습니다.

