



2010



**Академическая
программа SAS в России**

**THE
POWER
TO KNOW®**

Содержание:

1. Историческая справка о сотрудничестве SAS с Высшими школами в мире	3
В России и других странах СНГ участниками Академической Программы SAS являются:.....	3
2. Предложения SAS в области высшего образования:	4
А) Готовые курсы.....	4
Б) Подготовка курсов своими силами	5
С) Обучение и поддержка преподавателей.....	7
Д) Поддержка студентов.....	10
3. Информация об основном программном обеспечении SAS для учебного процесса	12
Кратко о системе SAS.....	12
<i>Доставка информации в масштабе предприятия</i>	<i>12</i>
Основное программное обеспечение SAS для учебного процесса	13
<i>JMP - Ваш превосходный визуализатор данных.....</i>	<i>13</i>
<i>Enterprise Guide</i>	<i>14</i>
<i>Дружественный Point&Click-интерфейс, предоставляющий удобный доступ ко всей мощи серверов SAS</i>	<i>14</i>
<i>Enterprise Miner</i>	<i>16</i>
<i>Интегрированная среда для получения знаний о бизнесе</i>	<i>16</i>

SAS Institute Inc
109004, г. Москва,
ул. Станиславского, 21, стр. 1
+7 (495) 937 4151
+7 (495) 937 4155

Данная брошюра предназначена для руководителей, сотрудников и студентов Высших учебных заведений и раскрывает возможности участия в Академической Программе SAS. В брошюре кратко описаны ресурсы Программы (в том числе и бесплатные) для преподавателей и студентов, а также дана справка по программному обеспечению, сопровождающему учебный процесс.

1. Историческая справка о сотрудничестве SAS с Высшими школами в мире

Глобальная Академическая Программа SAS предназначена для профессоров, студентов, исследователей и руководителей ВУЗов с целью использования программного обеспечения SAS в обучении, преподавании и исследовательских процессах.

В своей Глобальной Академической Программе компания SAS остается верна своей глубокой приверженности к мировому академическому сообществу и готова сотрудничать с Вами для поддержки учебных и исследовательских целей.

Полный список релизов по успешным проектам Академической Программы SAS здесь: <http://www.sas.com/success/wse.html>

Табл. №1 Краткая информация по количеству Вузов, участников Программы в мире

СТРАНА	КОЛИЧЕСТВО ВУЗОВ – УЧАСТНИКОВ ПРОГРАММЫ
США	> 2 400
Англия	> 70
Германия	> 120
Франция	> 200
Малайзия	>30
Австралия	>100

Академическая Программа SAS развивается и поддерживается почти в каждой стране, где присутствует отделение компании SAS.

В России и других странах СНГ участниками Академической Программы SAS являются:

- Московский институт управления и информационных технологий (МИИТ) при Московском государственном университете путей и сообщений
- Государственный Университет Управления
- Университетская математическая школа
- Московский энергетический институт
- Санкт-Петербургская Медицинская Академия общественного здравоохранения
- Киевский Технический Институт
- Западно-Казахстанский Государственный
- Западно-Казахстанский Государственный Медицинский Университет им. Марата Оспанова

Какие успехи были достигнуты Вузами, использующими SAS, можно проследить, познакомившись с релизами на веб-странице <http://support.sas.com/learn/ap/research/>, несколько релизов переведены на русский язык и представлены на российском сайте по адресу: <http://www.sas.com/offices/europe/russia/software/academic/index.html>

Рис. №1 Образец свидетельства участника SAS Академической Программы в России



2. Предложения SAS в области высшего образования:

А) Готовые курсы

■ Готовый курс Advanced Business Analytics для программы MBA (новинка)

Advanced Business Analytics - новейшая программа для обучающихся менеджменту. Она рассчитана на 15 недель обучения и включает в себя учебные пособия и демонстрационный материал.

“Из общения с представителями компаний, сотрудничающих с нами, мы поняли, что существует потребность в специалистах, понимающих роль аналитики в процессе принятия решений,” - комментирует исполнительный директор SAS Jim Goodnight. “Таким навыкам не обучают в школах MBA, и мы хотим первыми изменить это”.

Курс создан, чтобы продолжить базовые курсы статистики, которым выпускники MBA уже должны быть обучены. “Мы не делаем Data miner –ов этим курсом,” сказала преподаватель SAS Catherine Truxillo, которая разработала курс Advanced Business Analytics course. “Мы пытаемся помочь будущим лидерам оценить важность принятия решений, основанных на научном подходе и аналитике”.

Преподаватели, которые будут вести курс, получат: конспекты на 15 недель лекций и сопровождающие слайды PowerPoint slides; наборы данных, учебные примеры и доступ к программному обеспечению point-and-click SAS® OnDemand for Academics, включая SAS® Enterprise Guide®, SAS® Enterprise Miner™ и SAS Forecast Server, без дополнительной платы.

Кроме того, у студентов будет доступ к веб-сайту ресурса, который предлагает вводные вебинары по прогнозированию, Data Mining и подготовке данных; книги и огромное количество современных, ре-

levantных статей для обязательного чтения; и ряд рекомендательных видео от клиентов для обучения студентов.

Для поддержки обучения, SAS ведёт недельный учебный курс Advanced Business Analytics для преподавателей MBA.

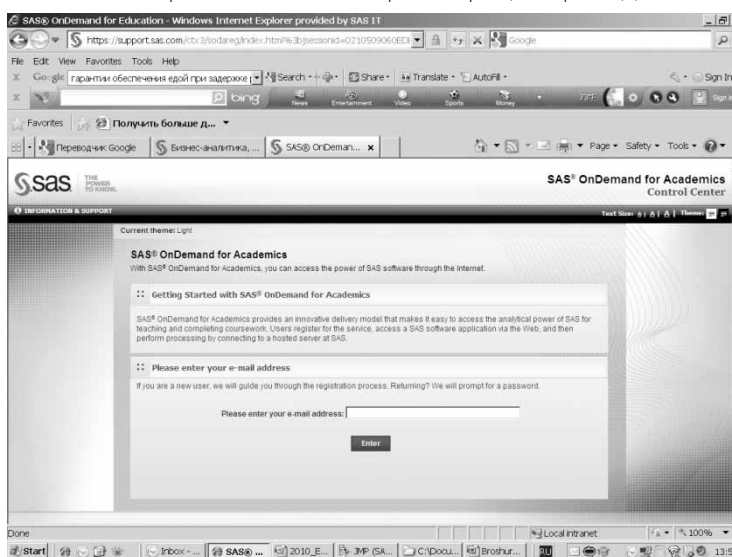
Б) Подготовка курсов своими силами

■ Программа On-demand Academy

http://www.sas.com/govedu/edu/programs/od_academics.html

Программа SODA (SAS On-Demand Academy) предоставляет on-line доступ для преподавателей и студентов к программным продуктам SAS Enterprise Guide, SAS Enterprise Miner (подробнее о свойствах программ в последнем пункте брошюры). Преподаватель может создать свой курс и загружать данные курса на сервер, используя FTP-доступ. Предусмотрена возможность просмотра результатов работы студентов, доступ которым дает преподаватель, используя SODA. Курс может быть использован на протяжении полугода. В нем заложена простая и наглядная навигация, point-and-click технология.

Рис. №2 Скрин-шот системы регистрации преподавателя



■ Использование программного обеспечения SAS и JMP

SAS объединила все важнейшие для использования в учебном процессе программы и объединила их в пакет SAS Education Analytical Suite (EAS). В нем собрано около 20 программ, которые позволяют строить курсы и проводить научные исследования. Курсы могут быть ориентированы как на уровень бакалавра, так и магистра или научных сотрудников. Сфера применения также широка - медицина, маркетинг, экономика, финансы, телекоммуникационные и информационные технологии, менеджмент - везде, где требуется изучение и анализ большого объема данных, технологии их извлечения, методологии анализа, построение деревьев решений и их трактовка.

Дополнительно предлагаются пакеты, предназначенные для узких целей отдельных Университетов, которые идут по специальной цене.

Отдельно стоит обратить на лицензии программного обеспечения JMP <http://www.jmp.com/academic/>, которое является превосходным визуализатором данных, снабженным более, чем 20-ю встроенными графическими визуализаторами (большая часть из которых трехмерна и с элементами демонстрации динамики по изменению данных) и более, чем 20-ю уже встроенными библиотеками моделей.

Лицензии на ПО продаются как для физических лиц (примерно 30 долл. за лицензию) прямо с сайта, так и для факультетов и кампусов. Многообразны варианты построения курсов и исследований, в которых возможно применение JMP, но наиболее часто его используют бизнес-школы, инженерные университеты и медицинские институты (для них существует специальный пакет – JMP Genomics).

■ Проведение мастер-классов от сотрудников российского офиса SAS или разработка заказных программ

Московское представительство SAS Russia имеет опыт организации и проведения мастер-классов силами ведущих российских специалистов, а также зарубежных коллег.

Мастер-классы могут проходить как бесплатно, так и на финансовой основе, в зависимости от тем и доступности специалистов.

Мастер-классы поддерживаются материалами раздела сайта SAS Press. Полезны также переведенные на русский язык издания: «Информационная революция» (Дж. Девис, Г. Миллер, А. Рассел), «Управление результативностью» (Г. Коккинз), «Аналитика как конкурентное преимущество» (Т. Дейвенпорт).

Компания SAS может помочь включить SAS в ваши академические программы. Если вы хотите создать новый курс и сертификационную программу, аналитические эксперты и консультанты SAS помогут разработать вам программу, необходимую для ваших целей. Для получения более подробной информации напишите, пожалуйста, по адресу: academic@sas.com.

■ Разработка сертификации

При накоплении собственной экспертизы Университет может заняться разработкой программы сертификации для студентов и преподавателей. SAS может выступить спонсором данного процесса и предоставить возможность использовать логотип SAS на итоговых сертификатах.

Табл. №2 Некоторые Университеты USA, разработавшие сертификационные программы при участии SAS

University of Central Florida	University of Alabama
Oklahoma State University	New York University
Central Michigan University	University of Louisville
College of Charleston	Kennesaw State University

С) Обучение и поддержка преподавателей

■ On-line documentations

Преподаватели могут бесплатно получить документацию по использованию программного обеспечения SAS.

Рис. №3 Скрин-шот сайта со страницы on-line-documentation



Для установки программного обеспечения пользователи следуют рекомендациям руководства по установке. При каких-либо затруднениях следует обращаться в службу поддержки московского отделения SAS по тел. +7 (495) 937 4151 .

■ E-learning

Участникам Академической программы поставляются вместе с лицензиями некоторые бесплатные электронные курсы. На остальные установлены скидки.

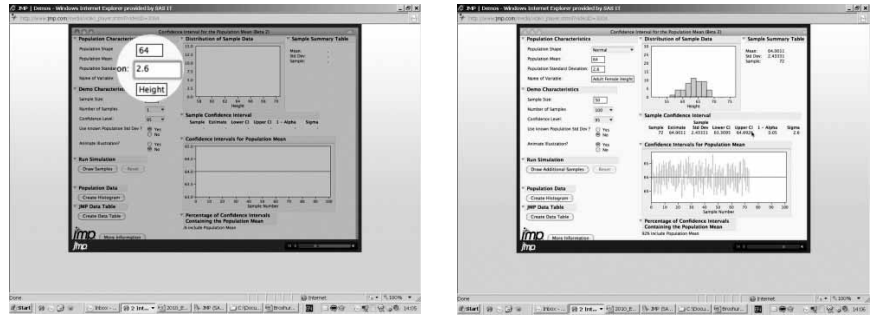
Есть открытые ресурсы по бесплатному начальному изучению SAS Enterprise Guide и JMP.

Бесплатные веб-касты (электронные семинары с обсуждением злободневных тем по аналитике или обучению SAS) можно найти по адресу: <http://www.sas.com/webcasts>.

Интернет-трансляции по различным технологиям и решениям от ведущих специалистов SAS и других компаний разных отраслей, можно найти по адресам:

Loyalty, Rewards and Value: What Do We Want from Our Customers?

Predictive Insurance Claims Processing: The Issue of Fraud



■ Очное обучение

Сотрудники ВУЗов, включенные в программу, имеют возможность обучаться на очных курсах Московского отделения SAS с 50% скидкой, а в некоторых случаях и бесплатно. Список тем - на сайте российского отделения SAS. Всем новым участникам Программы гарантирован бонус в виде бесплатного курса.

■ Учебные материалы

Преподаватели по запросу обеспечивают учебными материалами, которые содержат слайды, заметки к ним, установочные данные для курсов, концепции, нуждающиеся в обсуждении группы, практические примеры для студентов. Учебники поставляются в электронном виде на английском языке.

Материалы охватывают такие темы, как: Business Intelligence, Forecasting, Predictive Modeling , SAS

Programming, Statistics, Data Warehousing

■ Книги

<http://support.sas.com/publishing/authors>

SAS Press выпускает много книг по различным тематикам, которые могут быть полезны преподавателям в подготовке учебных программ и личного развития.

Категории книг: basic statistics, advanced analytics/predictive modeling, research/clinical, data mining, business intelligence

Участники программы с лицензией EAS имеют право на две бесплатные книги в год из списка SAS Press.

■ Бесплатные лицензии

Преподаватель Высшей школы, желающий построить курс с использованием SAS, но не имеющий возможности заключения договора между компаниями, имеет право прислать запрос на имя руководителя Учебного Центра SAS, и после заполнения некоторых документов, ему будет предоставлен бесплатный доступ к необходимой лицензии на полгода.

■ Летние школы

Ежегодное проведение бесплатных летних сессий для подготовки преподавателей. В 2010 г. планируются следующие темы:

По программированию

- Первый трек - New SAS teachers SAS® Programming 1 for High School
- Второй трек - Second Year Teachers SAS® Programming 2 for High School

По аналитике

- Advanced Business Analytics Course
- Data Mining for Educators Workshop

■ Новостная рассылка

Более 100 тыс. подписчиков по всему миру получают информацию о новостях Глобальной Академической Программы SAS по таким темам, как:

- SAS BUSINESS REPORT
- SAS TECH REPORT
- SAS STATISTICS AND OPERATIONS RESEARCH NEWS
- SAS RISK NEWS
- SAS FINANCIAL SERVICES NEWS
- SAS HEALTH AND LIFE SCIENCES NEWS
- SAS MANUFACTURING NEWS
- SAS RETAIL NEWS
- SAS TRAINING REPORT
- SAS PUBLISHING NEWS
- SAS CERTIFICATION NEWS

■ Сертификация преподавателей

Преподаватели имеют возможность получения ваучера на 50% скидку для прохождения профессиональной сертификации в тестовом центре Thompson Prometric. Ориентировочная рыночная стоимость теста – от 150 до 250 долл.

Рис. №5 Образец сертификата SAS



■ Академический клуб SAS в России

Клуб действует на базе Учебного центра в Москве. На его заседаниях рассматриваются интересные для преподавателей темы и проводятся презентации специалистами компании SAS. Электронная страница клуба: http://www.sas.com/offices/europe/russia/academic_club.

Рис. №6 Фотография с очередного заседания Клуба



■ Конференции и мероприятия

SAS ежегодно организует специальные мероприятия, посвященные аналитике, в которых на льготных условиях могут принимать участие представители Высших Школ. Так в 2010 г. прошли и планируются следующие мероприятия:

- SAS Global Forum (USA)
- F2010 Forecasting Conference (USA)
- M2010 Data Mining Conference (USA)
- A2010 Analytics Conference (Denmark)
- SAS Forum Russia (RF)
- The Premier Business Leadership Series (Germany, Hong Kong, USA)

Д) Поддержка студентов

■ Student Ambassador Program

Ежегодно 10 победителей программы отправляются бесплатно на SAS Global Forum с оплатой проезда, проживания и участия в различных мероприятиях.

■ Летние практики

Студенты российских Вузов имеют возможность прохождения практики при Московском представительстве SAS, изучая углубленно программное обеспечение SAS и принимая участие в проектах компании.

■ Академические дни

Академические дни проводятся для студентов на базе Вузов. ВУЗ отвечает за предоставление помещения и регистрацию студентов, а SAS - за анонс на сайте, приглашение лекторов, в том числе и из компаний, где используется SAS.

■ Point- and - Click - SAS Enterprise Guide

Изучение SAS Enterprise Guide возможно самостоятельно, используя следующие интернет-ресурсы:

1. Демонстрация возможностей программы (деморелик)
2. On-line документация
3. бесплатные электронные курсы и пособия по изучению Enterprise Guide
4. Скидки на очное обучение в классе офиса SAS
5. Доступ к данным для примеров и упражнений Enterprise Guide
6. On-demand Academy – on-line занятия на курсе, который организует преподаватель
7. Книга для самостоятельного изучения программы The Little SAS® Book for Enterprise Guide 4.1.

■ Point-and-Click – JMP для исследований в области статистики

Лицензию для студентов на JMP можно приобрести в режиме on-line. Стоимость лицензий на полгода – 29 долл. Там же есть бесплатные электронные курсы, лекции и вебинары с демонстрацией примеров по использованию статистических моделей.

■ Аспирантам и дипломникам

Аспиранты и дипломники, использующие в своей работе ПО SAS, могут продемонстрировать результаты своих работ специалистам SAS. Во время подготовки таких работ также возможна бесплатная консультация со стороны специалистов московского офиса.

■ Международное сообщество

По адресу SAS Student Network можно получить информацию о скидках на тренинги SAS, о том, как найти работу или пройти практику с использованием знаний о технологиях SAS.

3. Информация об основном программном обеспечении SAS для учебного процесса

Кратко о системе SAS

Доставка информации в масштабе предприятия

SAS System представляет собой интегрированную программную среду для доставки информации в масштабе предприятия. Функциональные возможности системы обеспечивают эффективное решение задач на каждом из четырех основных этапов процесса преобразования «сырых» данных в полезную информацию:

- Доступ к данным
- Преобразование данных и управление ими
- Анализ данных
- Представление данных

Программное обеспечение SAS используется для обработки преимущественно числовой информации и поддерживает работу с большими объемами данных.

SAS System может успешно применяться для решения задач в следующих областях:

- создание информационных систем для руководителей и систем поддержки принятия решений;
- ввод данных, извлечение данных из других систем и преобразование данных;
- организация хранилищ данных;
- создание отчетов и графиков;
- статистический и математический анализ;
- планирование и прогнозирование;
- исследование операций и управление проектами;
- статистический контроль качества;
- оценка производительности компьютерного оборудования;
- разработка приложений.

SAS System имеет модульную архитектуру, что позволяет организациям гибко подходить к использованию SAS System и лицензировать только те продукты, в функциональности которых они нуждаются. По мере того, как повышаются требования компании к системам доставки информации или изменяются ожидания пользователей в отношении предоставляемой функциональности, к уже лицензированным продуктам безболезненно могут быть добавлены другие компоненты. Центральным продуктом, обязательно присутствующим в любой конфигурации SAS System, является продукт Base SAS.

Программное обеспечение SAS может работать на различных аппаратных платформах, в различных операционных системах и поддерживает различные модели клиент-серверной архитектуры.

JMP – Ваш превосходный визуализатор данных

Вместо ограниченных и скучных статистических таблиц, полезных только для ответов на простые запросы и передачи обрывков информации, графический анализ JMP позволяет вам взаимодействовать с вашими данными в полном объеме для изучения взаимосвязей, определения тенденций, глубокого исследования интересующих областей, а также для открытия новых направлений, ранее не рассматривавшихся.

Первыми пользователями JMP были в основном ученые и инженеры, которые применяли программу для визуализации данных. С тех пор прошло 18 лет. Сегодня потребители используют более удобную функциональность JMP и динамическую визуализацию данных.

JMP может использоваться как самостоятельная программа в рамках отдельной рабочей станции, так и в качестве компонента платформы SAS Enterprise Intelligence. Все, кому необходимо принимать интеллектуальное решение на основе данных о бизнесе, найдут свое применение JMP. Пользователю не надо быть статистиком для получения информации на основе ваших данных. С JMP в этом просто нет необходимости.

Визуальное, интерактивное и легкое в освоении программное обеспечение JMP позволяет студентам быстро использовать мастер статистических концепций. Уникальная парадигма навигации JMP ведет студента по статистическим данным, используя технологии drag and drop, а затем отображает результаты в виде статистики и графиков в том же окне. Студенты могут легко выделять подмножества для лучшего понимания результата, с последующим анализом и созданием 3D и движущейся графикой. JMP работает на платформах Windows, Macintosh, Linux и легко импортирует данные из различных форматов, в т.ч. из Excel, SAS.

Рис. №7 JMP модели и графики



Дружественный Point&Click-интерфейс, предоставляющий удобный доступ ко всей мощи серверов SAS

Enterprise Guide – клиентское приложение для PC, которое переносит на настольные компьютеры всю мощь системы SAS. Это достигается за счет поддержки прозрачного доступа к данным посредством интерактивных диалогов, программ-экспертов и списков свойств, позволяющих легко выполнить все операции, необходимые для решения аналитических задач и подготовки отчетов любой степени сложности.

Прозрачный доступ к данным

- Доступ к данным SAS и к данным других источников, включая Microsoft Excel и Microsoft Access.
- Удаленный доступ к данным SAS на серверах версии 8, доступ к другим реляционным базам данных, включая Oracle и DB2.
- Обмен данными между локальным компьютером и сервером SAS при помощи операций «Drag&Drop»
- (например, перенос файла Microsoft Excel в таблицу на мэйн-фрейме или передача набора данных SAS с UNIX-сервера на компьютер пользователя).

Современный интерфейс

- Интерфейс полностью разработан с учетом стандартов Windows
- Проектно-ориентированный.
- Предоставляет Point&Click-интерфейс к языку SAS при помощи простых и удобных диалогов и «мастеров».
- Для более опытных пользователей системы SAS доступен новый редактор SAS-кода с цветовой подсветкой синтаксиса.
- Исчерпывающая сопроводительная документация.
- Настраиваемый пользовательский интерфейс:
 - › Полностью настраивается как на уровне пользователя, так и на уровне системы в целом.
 - › Меню и панели инструментов могут быть переопределены для удобства пользователя.



Свободный экспорт данных в другие приложения

- Полученные результаты могут быть:
 - › скопированы в буфер обмена и перенесены в другие приложения, такие как Microsoft Word или Microsoft Powerpoint;
 - › направлены в вашу систему электронной почты или сохранены в формате HTML для последующей публикации на Web-сервере;
 - › распечатаны и/или модифицированы для представления в требуемой форме отчетов.

Автоматизация

- Запуск проекта или любого его элемента может быть назначен на любое время в будущем на любом из имеющихся серверов SAS.
- Enterprise Guide обеспечивает непрерывную автоматическую публикацию обновленных данных на Web-сервер.

Поддерживаемые платформы

Windows 95, Windows 98, Windows NT V.4 и выше, Windows 2000

Поддерживаемые типы данных

- Табличные данные
 - › Данные, доступные через SAS:
 - наборы данных SAS (V6 – V8)
 - таблицы других СУБД (DB2 – Oracle)
 - › Другие источники:
 - Excel-Access-DB2-Lotus-Paradox-Text-HTML
 - ODBC-OLE/DB-MS Exchange
- Многомерные данные
 - › Данные SAS и другие источники:
 - OLE/DB for OLAP

Совместимость с ERP-системами

Enterprise Guide 1.2 сертифицирована SAP Portals как партнерская система по разработке программного обеспечения для доступа к SAP BW. Она предоставляет следующие возможности доступа к SAP BW:

- вывод на экран модели данных QueryCubes, определенных в SAP BW
- генерацию и формирование видов строк запросов MDX для извлечения данных из QueryCubes
- создание табличных и графических отчетов по QueryCubes

Enterprise Guide может функционировать как OLE DB для пользователя OLAP. В этом случае доступ из нее к данным SAP BW обеспечивается при помощи интерфейса ODBO API.

Интегрированная среда для получения знаний о бизнесе

Программный продукт Enterprise Miner – это интегрированный компонент системы SAS, созданный для выявления в огромных массивах данных информации, необходимой для принятия решений. Разработанный специально для поиска и анализа скрытых закономерностей в данных (data mining) Enterprise Miner включает в себя эффективные методы статистического анализа, соответствующую методологию выполнения проектов data mining и удобный графический интерфейс пользователя. Разработка проектов data mining может выполняться как локально, так и в архитектуре клиент-сервер. Важной особенностью Enterprise Miner является его полная интеграция с программным продуктом SAS/Warehouse Administrator, предназначенным для разработки и эксплуатации информационных хранилищ, и другими компонентами системы SAS. В частности, продукт SAS/IntrNet может эффективно использоваться для распространения результатов проекта data mining через локальную сеть или Интернет.

Графический интерфейс пользователя

Графический интерфейс является основным средством реализации проекта data mining. В его основу положен принцип “указать и щелкнуть”. Этот подход позволяет пользователю осуществить все стадии проекта data mining - от выбора источников данных до получения результатов.

Пользователь выбирает нужные функциональные узлы из палитры элементов и собирает из них блок-схему проекта. Данный графический интерфейс был разработан для двух групп пользователей. Бизнес-аналитики, имеющие минимальный опыт в области статистики, могут легко создавать собственные проекты из готовых самонастраиваемых модулей. Экспертам в области численного анализа предоставляется возможность более тонкой настройки аналитических процессов.

Методология SEMMA

Наличие соответствующей методологии разработки проектов data mining является ключом к эффективному использованию Enterprise Miner. Аббревиатура SEMMA образована от слов Sample (отбор данных), Explore (исследование отношений в данных), Modify (модификация данных), Model (моделирование взаимозависимостей), Assess (оценка полученных моделей и результатов). Методология SEMMA не предлагает разработчику проекта жестких правил, но предоставляет в его распоряжение научные методы построения концепции проекта, его реализации и оценки результатов проектирования.



Доступ к данным

Система SAS предоставляет пользователю доступ к данным, находящимся практически в любом формате. Для подключения к источнику данных и извлечения из него необходимого объема информации Enterprise Miner имеет три соответствующих функциональных модуля. На этой стадии проекта пользователь должен определить, как данные будут использоваться моделирующими узлами.

Исследование данных

Для визуального исследования данных имеются два графических модуля. Один из них предназначен для построения многомерных гистограмм и исследования распределений в больших объемах данных. Другой модуль используется для быстрого исследования зависимостей целевых переменных от входных параметров. Результат представляется в виде гистограмм и диаграмм рассеяния для входных и зависимых переменных.

На стадии анализа данных весьма эффективным инструментом является встроенный в Enterprise Miner продукт SAS/INSIGHT. С его помощью аналитик может проводить статистические исследования данных, причем интерактивно посредством множества диаграмм, взаимодействующих между собой.

Модификация данных

Для модификации данных в распоряжении аналитика имеется несколько функциональных узлов. С их помощью можно нормализовать данные, выделить нетипичные случаи в отдельное множество, заместить отсутствующие значения наиболее вероятными или по некоторому алгоритму.

Кластерный анализ

Цель кластерного анализа – поместить отдельные события в группы (или кластеры), причем так, что события, отнесенные к конкретному кластеру, имели наибольшее подобие между собой. В частности, с помощью кластерного анализа можно выявить группы покупателей в зависимости от типов продуктов, которые они приобретают. Для выполнения кластерного анализа в Enterprise Miner имеется специализированный модуль, позволяющий выполнять все виды кластерного анализа и определять автоматически оптимальное количество кластеров. Выявленные идентификаторы кластеров могут использоваться в дальнейшем анализе.

Ассоциации и последовательности

Модуль поиска ассоциаций позволяет выявлять в данных, как ассоциативные связи, так и последовательности и представлять их в виде правил. Работа данного модуля основана на подсчете условных вероятностей. Методы поиска ассоциаций и последовательностей широко используются для анализа покупательского спроса. В транспортной индустрии с помощью данного модуля, можно количественно оценить ассоциативную связь между совместными событиями; например, между перевозками нефтепродуктов.

Методы выявления последовательностей используются в том случае, если между событиями имеется временной промежуток. В этом случае, результатом анализа может быть совокупность правил следующего вида: «Из тех покупателей, которые воспользовались данной услугой, 25 процентов приобретут новую услугу в следующем месяце».

Дерево

Модуль «Дерево» необходим для автоматического построения классификационных и регрессионных деревьев решений. В первом случае для вывода правил сегментирования данных в распоряжение аналитика предоставлены критерии хи-квадрат и уменьшения энтропии. В случае регрессионных деревьев могут использоваться F-тест или критерий уменьшения дисперсии. Как и в случае кластерного анализа, данный модуль может использоваться для объединения отдельных событий в группы событий со схожими характеристиками, но с учетом причинно-следственных отношений.

Далее полученная модель дерева решений может использоваться для прогнозирования значений целевой переменной или для классификации событий. Например, при планировании маркетинговой кампании деревья могут быть эффективным инструментом для построения профилей покупателей, наиболее реагирующих на рассылку рекламной информации по почте. Во многих случаях методы деревьев решений имеют преимущество по сравнению с другими методами моделирования, так полученная модель может быть легко представлена графически и в виде совокупности правил.

Нейронная сеть

Методы нейронных сетей возникли в результате попыток моделирования нейрофизиологических процессов, происходящих в мозгу человека при распознавании образов. Модуль «Нейронная сеть» использует в своей работе алгоритмы регрессионных и дискриминантных моделей, объединенных в единую нелинейную динамическую систему. Нейронные сети используются для выявления нелинейных

отношений в данных, и, следовательно, для решения широкого класса проблем прогнозирования. В зависимости от задачи могут использоваться многослойный перцептрон, несколько разновидностей радиальных базисных функций или обобщенная линейная модель.

Регрессия

Регрессионный модуль предоставляет аналитику возможность выявления в данных линейных и логических зависимостей между целевой переменной и совокупностью независимых (входных) переменных. В случае линейной регрессии эти отношения выражаются в виде линейной функции целевой переменной от значений входных переменных. В случае логической регрессии будет построена функция для вычисления вероятности того, что целевая переменная будет равна некоторому дискретному значению. Для моделирования нелинейных регрессионных зависимостей может использоваться обобщенная линейная модель. Модуль включает в себя “Построитель взаимодействий”, позволяющий образовывать новые входные переменные более высоких порядков. Выбор значимых входных переменных может быть осуществлен автоматически одним из трех методов: прямым, обратным и ступенчатым.

Оценка результатов моделирования и вычислений

Каждый модуль производит необходимую диагностическую информацию необходимую для проверки качества полученных результатов. Для выявления наиболее адекватных моделей и организации обработки новых данных (в том числе и для вычисления прогнозов) имеются соответствующие модули. Еще один модуль предоставляет разработчику возможность автоматического документирования всего проекта data mining в формате HTML.



**THE
POWER
TO KNOW.**

Московское представительство SAS Institute GmbH:

109004 г. Москва, ул. Станиславского, дом 21, строение 1, • Тел: +7 495 937 4151 • Факс: +7 495 937 4155 • www.sas.com/russia

Представництво «САС ІНСТІТЮТ ЧР, С.Р.О.» Україна,

01601, Київ, вул. Шовковична 42-44 • Тел: +38 (044) 459 0355 • Факс: +38 (044) 490 1200

SAS Institute Inc. Мировая штаб-квартира

SAS Campus Drive, Cary, NC 27513 USA. • Тел: +1 919 677 8000 • Факс: +1 919 677 4444 • <http://www.sas.com>

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration. Other brand and product names are trademarks of their respective companies. Copyright © 2006, SAS Institute Inc. All rights reserved.