

Mit der beeindruckenden Palette an Data-Mining-Algorithmen des *Data Miner* decken Studenten der Oklahoma State University unbekannte Muster auf, erklären diese und erstellen Prognosen

Hintergrund. Die Oklahoma State University (OSU) ist eine bekannte fachübergreifende staatliche Universität und zugleich eine Forschungseinrichtung. Neben den vielen traditionellen Studenten sind auch viele nicht-traditionelle Studenten eingeschrieben, die Vollzeit in der Industrie arbeiten. Diese nicht-traditionellen Studenten bringen eine Vielzahl an Klassifikations- und Vorhersageproblemen ein, die mit den Werkzeugen und Techniken des *STATISTICA Data Miner* gelöst werden. Einige der letzten Aufgaben betrafen u. a. die Vorhersage von Zuckerkrankheiten auf Basis von demografischen, sozialen und die Freizeit betreffenden Parametern, Zielmarketing-Modelle für die Optimierung von Mailings und die Vorhersage finanzieller Indikatoren wie dem S&P 500 und von Umrechnungskursen.

***STATISTICA Data Miner* und die Konkurrenz.** Nach Betrachtung aller führenden Data-Mining-Tools wählte die Oklahoma State University basierend auf den folgenden Faktoren den *STATISTICA Data Miner*:

- Er besitzt eine beeindruckende Palette an Data-Mining-Algorithmen.
- Er hat eine interaktive grafische benutzerfreundliche Oberfläche, einschließlich einer Workflow-Oberfläche. Auf dieser kann mithilfe von grafischen Symbolen und Verknüpfungspfeilen ein vollständiges Projekt erstellt werden, dessen Verfahren alle gleichzeitig laufen können.
- Das Programm besitzt überragende grafische Möglichkeiten.
- Der *STATISTICA Data Miner* kann als Webanwendung gestartet werden, so dass jeder mit einer Berechtigung das Programm über einen Webbrowser nutzen kann, ohne Komponenten auf Client-Seite installieren zu müssen.

Dr. Dursun Delen, Assistenzprofessor im Fachbereich für Management und Informationssysteme: „Diese Merkmale haben den *STATISTICA Data Miner* zur ersten Wahl bei der Verwendung von Programmen im Unterricht über Data Mining bei diplomierten IT-Studenten und für unsere Forscher gemacht.“ Außerdem war der *STATISTICA Data Miner* trotz seiner Vorteile gegenüber anderen Programmen nicht teurer als andere umfangreiche Programmpakete auf dem Markt.

Die Stärken des *STATISTICA Data Miner*. Der *STATISTICA Data Miner* stellt alle Mittel, die Studenten an der OSU für ihre Data-Mining-Projekte benötigen, zur Verfügung. Dr. Delen: „Wir sind von den vielen Stärken des *STATISTICA Data Miner*, der umfangreichen Palette an Algorithmen, der großen Auswahl an Werkzeugen zur Visualisierung und der imponierenden Geschwindigkeit beeindruckt. Die Algorithmen betreffend kann ich basierend auf meiner jahrelangen Erfahrung in Industrie und Hochschule mit Sicherheit sagen, dass der *STATISTICA Data Miner* mit die umfangreichsten Data-Mining-Algorithmen auf dem Markt besitzt, die solche Anwendungsgebiete wie Vorhersagen (z. B. Regression), Klassifikation (z. B. Klassifikationsbäume, neuronale Netze, Diskriminanzanalyse), Clustering (z. B. k-Means, Kohonen SOM) und das Aufdecken von Assoziationsregeln (z. B. A Priori-Algorithmen) abdecken.“

Zusätzlich zu einem der effizientesten implementierten Modellierungs-Algorithmen sind auch eine große Auswahl an Algorithmen zur Aufbereitung und Visualisierung der Daten im *STATISTICA Data Miner* integriert. Das Produkt nutzt Multiprozessor-Umgebungen, indem es parallele Verarbeitung auf Server-Seite unterstützt, und dadurch die Geschwindigkeit bei der Durchführung von Data-Mining-Projekten für Studenten der OSU, die näherrückende Abgabetermine einzuhalten versuchen, erheblich erhöht. Das Programm stellt alle Arten von grafischen und tabellarischen Ausgaben zur Verfügung. Basierend auf der Auswahl des Anwenders in der Oberfläche zur Projektspezifikation kann jeder Data-Mining-Algorithmus einfache, erweiterte oder umfangreiche Berichte zur Ausgabe erstellen. Der *STATISTICA Data Miner* stellt Anwendern auch eine Deployment-Funktionalität zur Verfügung, mit der der Endanwender die entwickelten Data-Mining-Modelle leicht in Produktionssystemen einsetzen und verwerten kann.

Dokumentation und Support. Studenten der OSU halten die Produktdokumentation für umfangreich und das Schritt-für-Schritt-Lehrbuch für ein wertvolles Werkzeug, um das Programm schnell kennenzulernen.

Dr. Delen erinnert sich an die positive Erfahrung mit dem technischen Support von StatSoft während der Implementierung. „Wir haben von StatSoft herausragenden Support erhalten. Sie kamen zu uns, um die Software auf unserem Webserver zu installieren, und haben alle unsere Fragen zum Programm und zu seinen Merkmalen beantwortet. Sie waren während der gesamten Dauer des Prozesses sehr geduldig, zuvorkommend und hilfreich.“

