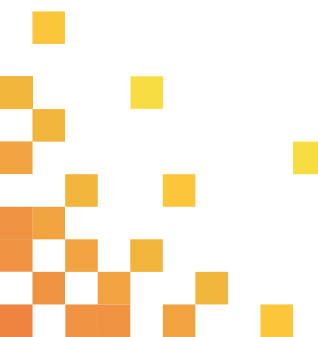
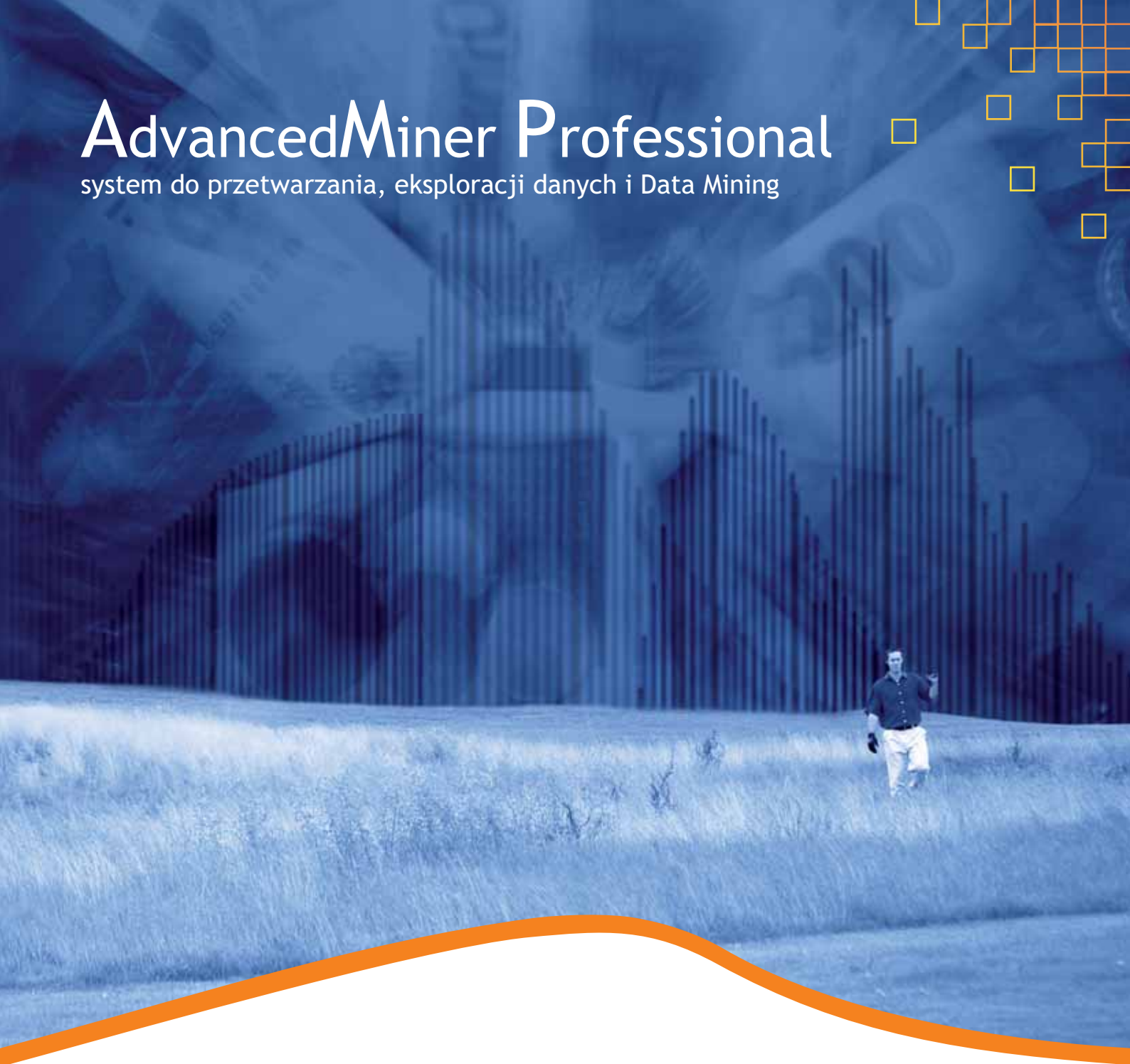


# AdvancedMiner Professional

system do przetwarzania, eksploracji danych i Data Mining



# AdvancedMiner Professional

## system do przetwarzania, eksploracji danych i Data Mining

**AdvancedMiner Professional** to nowoczesne i sprawdzone na rynku rozwiązanie umożliwiające kompleksową realizację różnorodnych projektów analitycznych.

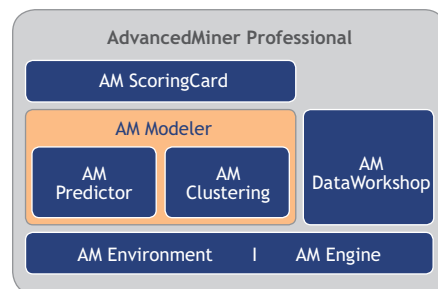
Mogą one dotyczyć na przykład:

- zwiększenia skuteczności kampanii marketingowych,
- obniżenia ryzyka przejścia klientów do konkurencji,
- oceny wiarygodności kredytowej klientów (Credit Scoring),
- tworzenia data martów na potrzeby analityczne lub raportowe,
- automatyzacji procesu weryfikacji jakości zgromadzonych danych.

Oferowane przez firmę StatConsulting rozwiązania obejmują nie tylko oprogramowanie, ale również szkolenia i odpowiednio zaprojektowane usługi zwiększające efektywność pracy z systemem.

### Warianty wdrożenia

- **AM Environment** - bazowe środowisko pracy analityka,
- **AM Engine** - wbudowany wydajny, bazodanowy silnik obliczeń,
- **AM DataWorkshop** - aplikacja umożliwiająca wszechstronną pracę z danymi i raportami,
- **AM Predictor** - rozwiązanie dedykowane projektom przewidywania i modelowania zachowań klientów,
- **AM Clustering** - narzędzie umożliwiające efektywną realizację projektów segmentacji danych,
- **AM Modeler** - pakiet obejmujący AM Predictor i AM Clustering,
- **AM ScoringCard** - narzędzia do konstrukcji i walidacji kart scoringowych, analiz ad-hoc wykonywanych przy projektach Credit Scoring i Credit Rating.

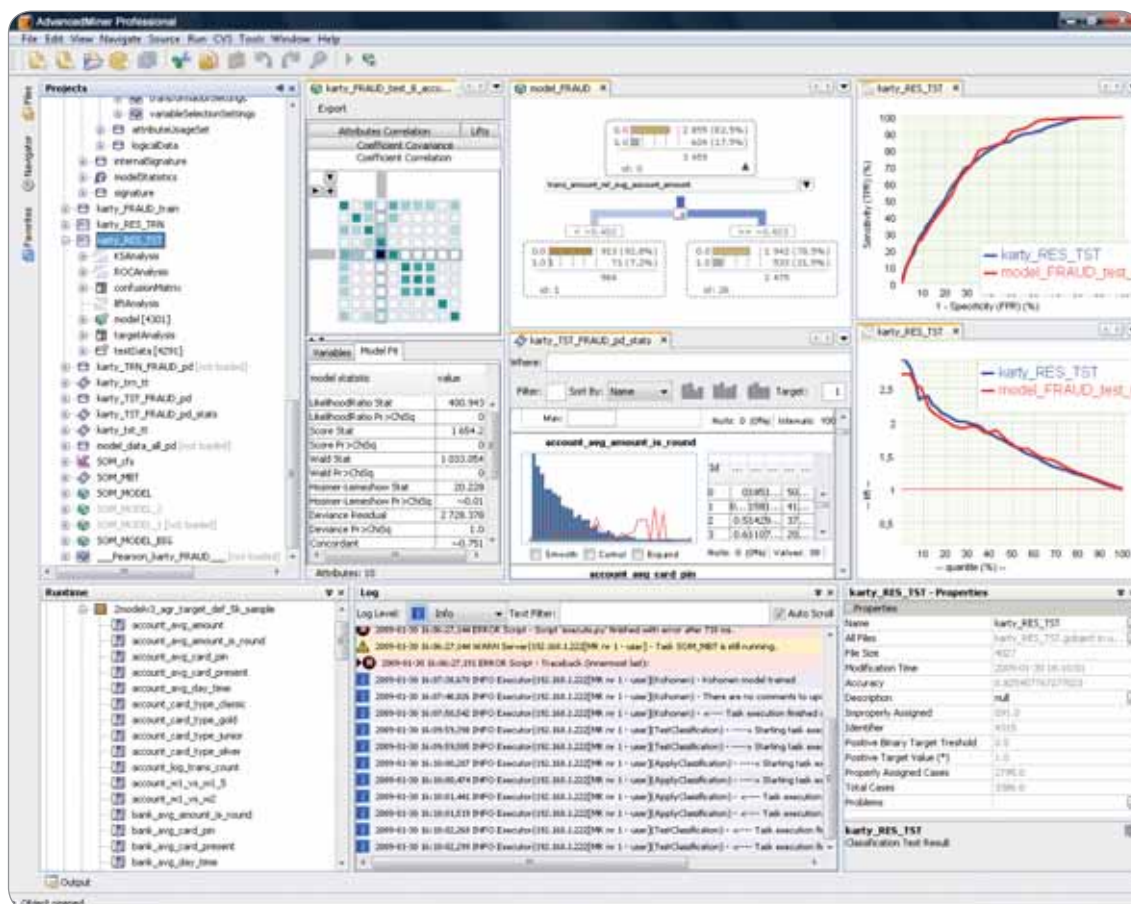


## Ogólna charakterystyka systemu

*AdvancedMiner Professional* to kompletne i zintegrowane środowisko do prowadzenia projektów analitycznych, które oferuje różnorodne narzędzia wspierające pracę analityka i programisty. Nasze oprogramowanie nie tylko umożliwia przetwarzanie danych ale i udostępnia szereg algorytmów statystycznych i Data Mining pozwalających na budowanie efektywnych modeli analitycznych.

System powstał z myślą zarówno o początkujących jak i zaawansowanych analitykach. Początkujący użytkownicy mogą zbudować pierwsze modele i eksperymentować z ustawieniami algorytmów, korzystając z łatwego w obsłudze interfejsu graficznego i ustawień domyślnych. Zaawansowani analitycy i eksperci odkryją wartościowe narzędzie analityczne, które można modyfikować i rozbudowywać. Jest to możliwe, między innymi, dzięki językowi skryptowemu - Gython, który bazuje na języku Python i jest w pełni funkcjonalnym językiem programowania rozszerzonym o funkcje typowe dla analizy danych.

Nasz system pozwala także na pracę zespołową: analitycy mogą dzielić się obiektami poprzez wspólne repozytoria metadanych lub lokalne i zdalne katalogi. Zarządzanie skryptami i bibliotekami utworzonymi przez użytkowników odbywa się za pomocą systemu kontroli wersji CVS, a przetwarzane dane mogą być współdzielone. Jest to narzędzie, które może pracować na większości popularnych systemów operacyjnych i korzystać z różnych systemów bazodanowych.



Graficzne środowisko systemu *AdvancedMiner Professional*

## Narzędzia systemu *AdvancedMiner Professional*

### Obsługa zbiorów danych

Oferowany przez StatConsulting system może być z powodzeniem stosowany przez firmy, których specyfika działalności wymaga gromadzenia dużych wolumenów danych (np. telekomunikacja, bankowość).

### *AdvancedMiner Professional* umożliwia wczytywanie danych z:

- praktycznie dowolnych systemów relacyjnych baz danych wspierających standard ODBC/JDBC (m.in. MS SQL, MySQL, Oracle, Sybase),
- hurtowni danych,
- plików tekstowych w standardzie CSV,
- arkuszy kalkulacyjnych.

Użytkownik może przeglądać zawartość bazy danych, a także ją edytować i tworzyć nowe tabele, dzięki wbudowanemu językowi skryptowemu. Umożliwia on korzystanie z instrukcji SQL oraz udostępnia specjalne konstrukcje językowe przeznaczone do przetwarzania danych.

The screenshot displays the AdvancedMiner Professional interface. On the left, a tree view shows the 'My\_Database' structure with various tables like 'german\_credit'. The main window shows a SQL query being executed on the 'german\_credit' table. The query uses aggregate functions (MIN, AVG, MAX) to calculate statistics for different personal statuses and housing types. Below the query, the results are displayed in a table format.

[Index]	personal_status	housing	min_duration	avg_duration	max_duration	avg_age	avg_credit_amount
0	female div/dep/mar	for free	6	22.474	60	48.105	5 165.368
1	female div/dep/mar	own	4	19.204	60	33.669	2 693.556
2	female div/dep/mar	rent	4	19.316	48	27.958	2 820.958
3	male div/sep	for free	12	20.0	24	42.667	4 772.667
4	male div/sep	own	6	21.512	48	38.171	3 778.073
5	male div/sep	rent	6	20.0	36	39.833	2 502.667
6	male mar/wid	for free	24	24.0	24	47.0	5 103.0
7	male mar/wid	own	6	18.147	72	31.25	1 854.162
8	male mar/wid	rent	6	16.13	42	26.957	2 478.435
9	male single	for free	6	28.871	60	42.859	4 850.682
10	male single	own	4	21.113	60	36.993	3 371.297
11	male single	rent	6	20.327	60	34.927	3 980.473

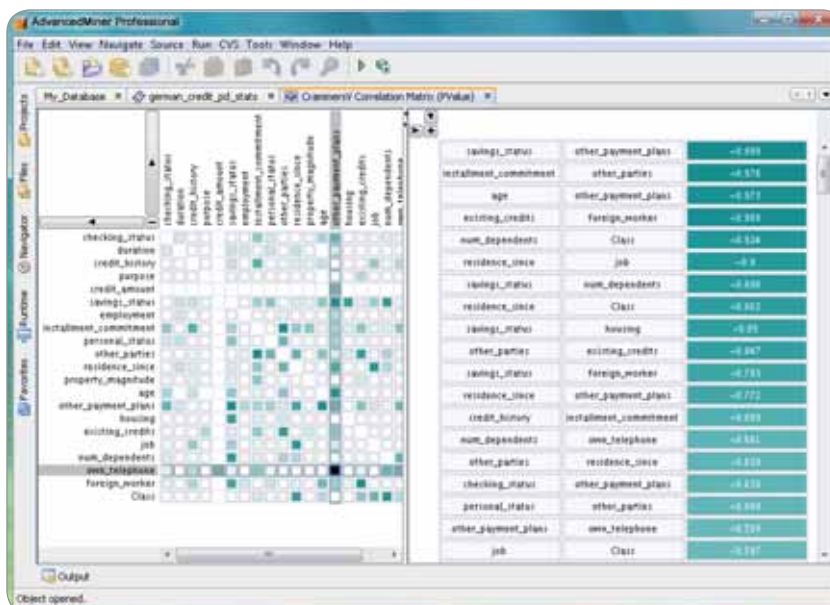
### Praca z bazami danych

*AdvancedMiner Professional* jest wyposażony również w szereg standardowych transformacji (np. binaryzacja, standaryzacja) przydatnych podczas analizy danych, które oszczędzają czas i ułatwiają pracę analityka. Ponadto, istnieje możliwość definiowania własnych transformacji przez użytkownika.

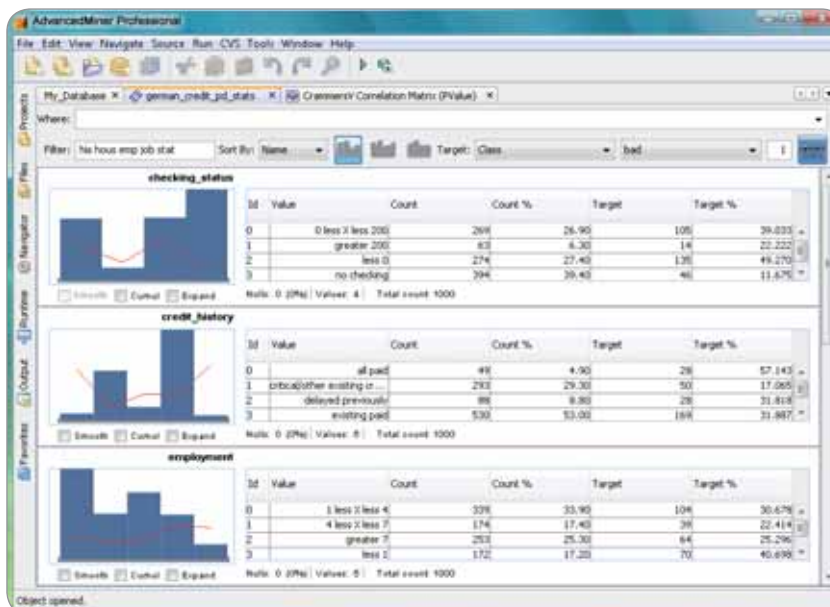
## Obsługa zmiennych

Wraz z systemem *AdvancedMiner Professional* użytkownik uzyskuje dostęp do szerokiego zakresu funkcjonalności kompleksowo wspierającego go w zakresie eksploracji danych. System wyposażony jest zarówno w standardowe procedury statystyczne jak i wyspecjalizowane narzędzie do interaktywnej graficznej eksploracji danych. Umożliwia ono między innymi:

- wizualizację rozkładów,
- określanie zależności ze zmienną objaśnianą (target),
- tworzenie zmiennych wirtualnych wspomagających analizy,
- grupowanie wartości zmiennych,
- analizę korelacji zmiennych (m.in. Pearson, Correlation Ratio, ChiSquare).



Macierz korelacji



Zmienne katetyczne

## Modelowanie

**Advanced Miner Professional** posiada narzędzia do rozwiązywania różnych zadań z zakresu:

- klasyfikacji,
- aproksymacji,
- grupowania,
- odkrywania reguł asocjacyjnych,
- analizy czasu przetrwania.

Do konstrukcji modeli analitycznych mogą być wykorzystane algorytmy, takie jak między innymi:

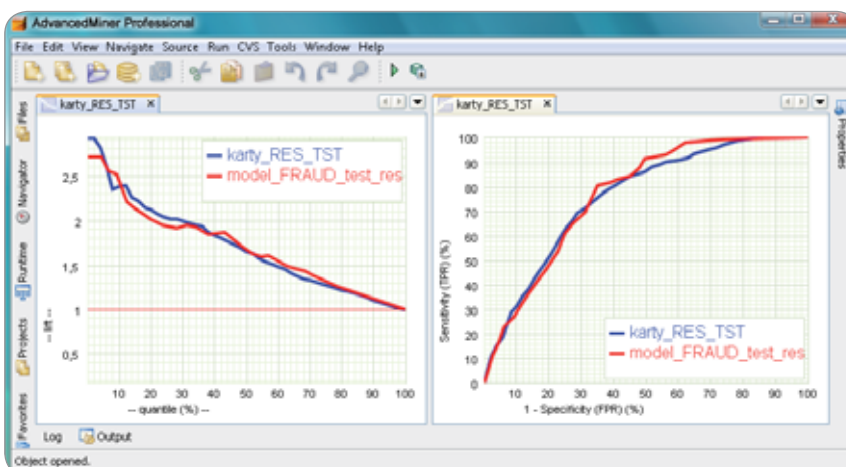
- drzewa klasyfikacyjne,
- regresja liniowa,
- regresja logistyczna,
- k-means,
- reguły asocjacyjne.



Drzewo klasyfikacyjne

**Advanced Miner Professional** umożliwia wszechstronną ocenę oraz porównywanie jakości modeli, przy zastosowaniu takich narzędzi jak LIFT, ROC, K-S, Confusion Matrix. Cały proces odbywa się w łatwym w obsłudze, interaktywnym środowisku graficznym.

Raporty dotyczące jakości modeli możliwe są do wygenerowania w formacie MS Office lub OpenOffice.



Ocena jakości modeli Data Mining

## Kod scoringowy

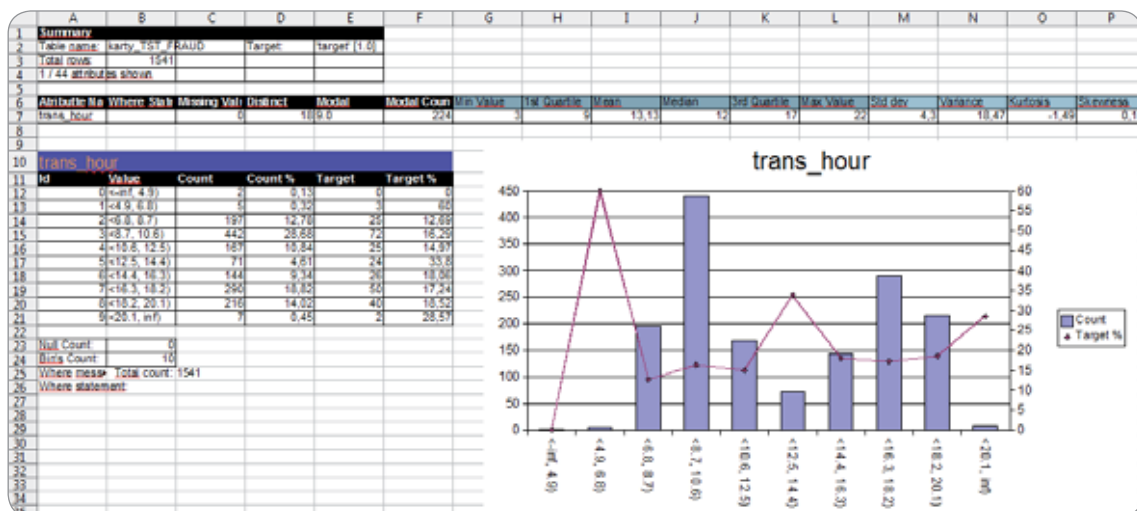
System umożliwia generowanie kodu scoringowego (w różnych formatach), co pozwala na korzystanie z utworzonych modeli niezależnie od systemu *AdvancedMiner Professional*.

Dzięki takiemu rozwiązaniu, można znacznie uprościć i przyspieszyć wdrożenie modeli poprzez łatwą integrację z innymi systemami informatycznymi firmy.

## Raportowanie

W systemie *AdvancedMiner Professional* możliwe jest automatyczne i półautomatyczne tworzenie cyklicznych raportów zawierających szeroki zakres zestawień i statystyk bez konieczności poświęcania znacznej ilości czasu na ich przygotowanie. *AdvancedMiner Professional* umożliwia tworzenie skryptów budujących raporty wykorzystujące między innymi: sformatowany tekst, pola autotekstu, wyniki obliczeń, tabele, wykresy.

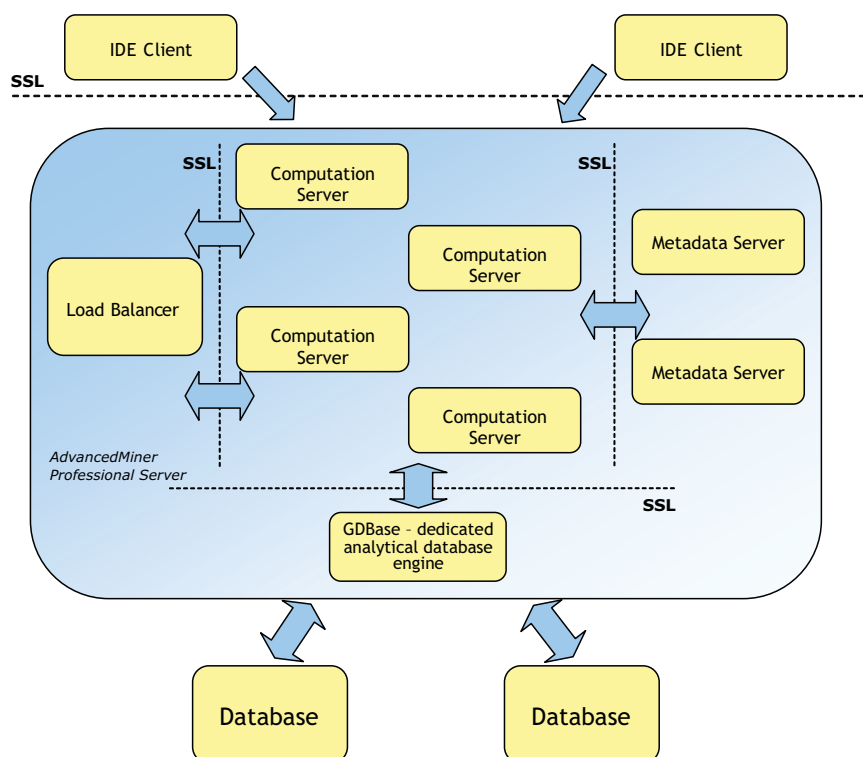
Raport może być wyeksportowany do pakietów biurowych MS Office, OpenOffice, w postaci dokumentu PDF oraz dokumentu HTML.



Przykład fragmentu raportu wygenerowanego w systemie *AdvancedMiner Professional*

## Architektura systemu

System *AdvancedMiner Professional* wykorzystuje architekturę Client - Server, przy czym serwer fizycznie może zostać rozproszony na wiele jednostek obliczeniowych. Komunikacja między jednostkami obliczeniowymi odbywa się za pomocą protokołu SSL (Secure Socket Layer). Daje to analitykowi możliwość bezpiecznej pracy wielozadaniowej. Oczekując na wyniki zdefiniowanego zadania, analityk może zająć się przygotowaniem kolejnego, a także rozpocząć jego wykonywanie na tym samym lub innym serwerze. Przyspiesza to proces analizy danych, zapewniając jednocześnie pełne wykorzystanie dostępnych zasobów. Dzięki temu, możliwe jest stworzenie wydajnego i skalowalnego systemu do przetwarzania i analizy danych, wykorzystując tańsze komputery klasy PC nawet z różnymi systemami operacyjnymi.



Schemat architektury systemu *AdvancedMiner Professional*

### Kompatybilność

*AdvancedMiner Professional* oparty został na sprawdzonych technologiach ze środowiska Java, zapewniających niezależność od platformy. Pracuje on zarówno pod systemami MS Windows jak również pod systemami z rodziny Unix (w tym systemy operacyjne Linux).

System współpracuje z relacyjnymi bazami danych, które udostępniają interfejs JDBC/ODBC (np. MySQL, MS SQL, Oracle, Sybase SQL Anywhere Studio). Ponadto, użytkownik ma możliwość skorzystania z bazy wewnętrznej systemu - GDBase.

## Zalety systemu *AdvancedMiner Professional*

### ■ Integracja z zewnętrznymi systemami informatycznymi

- wczytywanie danych z różnych systemów relacyjnych baz danych (MS SQL, MySQL, Oracle, Sybase oraz innych wspierających standard ODBC/JDBC), hurtowni danych, plików tekstowych oraz z arkuszy kalkulacyjnych,
- generowanie kodu scoringowego, który może obejmować zarówno modele jak i transformacje danych,
- raportowanie zintegrowane z pakietami biurowymi MS Office oraz OpenOffice,
- kompatybilność zarówno z systemami MS Windows, jak również systemami z rodziny Unix.

### ■ Rozproszenie obliczeń

- możliwość fizycznego rozproszenia serwera na wiele jednostek obliczeniowych, co pozwala na przyspieszenie procesu analizy danych przy pełnym wykorzystaniu dostępnych zasobów,
- możliwość stworzenia wydajnego i skalowalnego systemu przy niskich kosztach, wykorzystując tańsze komputery klasy PC z różnymi systemami operacyjnymi (np. Linux, MS Windows).

### ■ Przetwarzanie danych

- brak ograniczeń na liczbę kolumn przetwarzanych danych,
- wysoka wydajność operacji bazodanowych w ramach zintegrowanej analitycznej bazy danych (GDBase),
- możliwość pracy z bardzo dużymi bazami danych (np. 5 mld rekordów) zarówno na dedykowanym serwerze systemu jak i stacjach roboczych klasy PC.

### ■ Zintegrowane środowisko modelowania

- pełna funkcjonalność w zakresie eksploracji danych dostępna w nowoczesnym, interaktywnym i graficznym narzędziu,
- bogate możliwości porównywania modeli w trakcie ich prototypowania i testowania,
- łatwe konstruowanie modeli hybrydowych, łączących kilka różnych metod analizy.

## Zalety systemu *AdvancedMiner Professional* c. d.

### ■ Szeroki wachlarz zastosowań

- np. przetwarzanie i eksploracja danych, Credit Scoring, przewidywania Churn, analizy wspierające Cross/Up selling, segmentacja klientów, analizy LTV i Survival.

### ■ Efektywna obsługa

- elastyczność systemu dzięki dostępności języka skryptowego pozwalającego na pracę z danymi, modelami,
- możliwość szybkiego dostosowania rozwiązania do indywidualnych potrzeb oraz konstrukcji nowych wyspecjalizowanych rozwiązań,
- dostępność zaawansowanego edytora wspierającego konstrukcję modeli o szerokim zakresie funkcjonalności niedostępnym w innych tego typu narzędziach,
- jednorodne podejście obiektowe umożliwiające efektywne używanie różnorodnych metod,
- wsparcie dla pracy zespołowej obejmujące CVS (Concurrent Version System), dzielenie się metadanymi i danymi.

Jeżeli są Państwo zainteresowani uzyskaniem szczegółowych informacji na temat systemu *AdvancedMiner Professional*, zachęcamy do skontaktowania się z nami. Z przyjemnością udzielimy Państwu dodatkowych informacji oraz odpowiemy na pytania.

## Charakterystyka firmy StatConsulting

Firma StatConsulting została założona w 2001 roku w Warszawie. Oferujemy wsparcie analityczne w zakresie zarządzania ryzykiem (kredytowym, rynkowym, operacyjnym w kontekście wymagań Nowej Umowy Kapitałowej i Solvency II) oraz w obszarze modelowania zachowań klientów (Customer Intelligence, analityczny CRM).

Naszą wiedzę wykorzystujemy do przeprowadzania analiz dla firm działających w branżach o wysokim poziomie konkurencji. Specjalizujemy się w świadczeniu usług dla sektora finansowego oraz telekomunikacyjnego. Przeprowadziliśmy wiele projektów analitycznych związanych z konstrukcją różnych rodzajów modeli (m.in. CRM analityczny, scoring). Posiadamy bogate doświadczenie w budowie, walidacji oraz monitoringu modeli scoringowych opartych zarówno na danych aplikacyjnych, jak i behawioralnych. Mamy praktyczne doświadczenie w budowie analitycznych repozytoriów danych, przygotowaniu danych, integracji z danymi zewnętrznymi, konstrukcji modeli oraz ich implementacji.

Wysoka jakość dostarczanych usług oraz bogate doświadczenie w implementacji i prowadzeniu projektów analitycznych pozwala nam utrzymać znaczącą pozycję wśród firm świadczących usługi z zakresu analizy danych.

## Wybrani klienci

### Sektor finansowy i bankowy

ALG Bank S.A.  
Alior Bank S.A.  
Bank BPH S.A.  
Bank Pocztowy S.A.  
Bank Ochrony Środowiska S.A.  
BRE Bank Hipoteczny S.A.  
BRE Leasing Sp. z o.o.  
Euro Bank S.A.  
Europejski Fundusz Leasingowy S.A.  
Gospodarczy Bank Wielkopolski S.A.  
Kruk S.A.  
Mercedes-Benz Bank Polska S.A.  
Mercedes-Benz Leasing Polska Sp. z o. o.  
PolCard S.A.  
PTE PZU S.A.  
PZU Życie S.A.  
Raiffeisen-Leasing Polska S.A.  
Toyota Bank Polska S.A.

### Sektor telekomunikacyjny

Polkomtel S.A.

### Sektor farmaceutyczny

Cegedim Polska Sp. z o.o.  
Eli Lilly Polska Sp. z o.o.

### Sprzedaż detaliczna oraz FMCG

Bertelsmann Media Sp. z o.o.  
Carrefour Sp. z o.o.  
MDS Poland Sp. z o.o.

### Przemysł oraz energetyka

KGHM Polska Miedź S.A.  
Klingspor Sp. z o.o.  
Statoil Polska Sp. z o.o.

### Marketing bezpośredni oraz agencje badania rynku

ASM - Centrum Badań  
i Analiz Rynku Sp. z o.o.  
Premium Club Sp. z o.o.



**STATCONSULTING**

StatConsulting Sp. z o.o.  
ul. Wołoszowskiego 38a  
02-724 Warszawa  
tel.: (+48) 22 / 847 97 17  
fax: (+48) 22 / 499 45 31  
e-mail: [info@statconsulting.com.pl](mailto:info@statconsulting.com.pl)  
[www.statconsulting.com.pl](http://www.statconsulting.com.pl)

