

my_X: Maßgeschneiderte online Analysen

László Pitlik

Institut für Wirtschaftsanalyse
Universität Gödöllő
Páter K. u. 1.
H-2100 Gödöllő,
pitlik@miau.gau.hu

Abstract: Das Ziel des staatlich geförderten Innovationsprojektes besteht darin, anhand von verhältnismäßig wenigen Daten für jede Person/Unternehmen/Institution maßgeschneiderte Analysen online anzubieten. Diese Dienstleistung ist eine Art Demokratisierung der Data-Mining-Philosophie, durch die man bislang anhand von großen Datenmengen, rel. teuren Hardware-Ausstattungen und Software-Lizenzen, sowie gut bezahlten Experten die Analysewünsche des Unternehmens ad hoc oder regelmäßig abdeckte.

1 Vorgeschichte des Vorhabens

An der Universität Gödöllő stehen seit Jahren im Zentrum des Wirtschaftsinformatik-Unterrichtes und der Forschung (OTKA T-049013) - als Kern der Entscheidungsunterstützung - die Anwendungsmöglichkeiten von solchen künstlichen Intelligenzmethoden, welche Studenten aus diverser Fachrichtungen im Rahmen von Tabellenkalkulationen nachvollziehen können. Die Studenten sollten zunächst die verhältnismäßig wenigen Inputdaten einer möglichst realistischen Entscheidungssituation als Grundlage von Pivot-Tabellen (bzw. immer öfter als OLAP-Dienste) bereitstellen. Aus diesen Rohdaten können solche Objekt-Attribut-Matrizen durch die pivot/olap-Unterstützung abgeleitet werden, welche dann direkt die Eingangsdaten des fallbasierten Schließens darstellen. Die angebotenen Lösungsansätze können den Expertensystemen, den neuronalen Netzen und den Ähnlichkeitsanalysen zugeordnet werden. Erfahrungen mit Hunderten von Studenten weisen darauf hin, dass die technische Realisierung zwar eingeübt werden kann, diese Phase nimmt jedoch unnötig viel Zeit davon ab, wie man Probleme erkennt und/oder Lösungen interpretiert. Die hohen Anzahl der Fallstudien und die breite Skala der erkannten Problembereiche haben überzeugende Impulse dazu geliefert, die ganze Analysetätigkeit im Rahmen von passenden online Diensten möglichst automatisiert anzubieten. Gut ausgebildete (für ihr Umfeld „infektionsfähige“) Kunden entstehen also Jahr für Jahr aus den eigenen Studenten. Im etwa selben Anteil (ca. 10%) als unter den Studenten sollten auch unter den alltäglichen Entscheidungsträgern solche Individuen existieren, die mögen, hoch strukturiert und faktenbasiert zu denken.

Daher wurde ein Antrag an das ungarische Amt für Innovation (NKTH) gestellt. Das Vorhaben (my_X: my Expertises) wurde positiv bewertet und nun ab Januar 2007 besteht die Möglichkeit, eine mehrsprachige Dienstleistung zusammenzustellen.

2 Darstellung der potentiellen Dienste

Während des Projektes soll ein benutzerfreundliches Rahmensystem entstehen, welche folgende Aufgaben hat: Registrierung, Zugangsregelung (Daten, Analysen), e-Learning-Tools (Lexikon, Online Bibliothek, Nachrichten, etc.) für Aus- und Weiterbildung der Kunden, OLAP-Simulationen, Statistiken, Online Zahlungen, etc. Diese notwendigen, jedoch nicht besonders innovativen Module werden nach den bekannten „Standards“ ähnlicher Systeme bereitgestellt.

Zu den sog. Kerndiensten zählen die numerischen (Ähnlichkeitsanalysen, Prognosen) und die logischen Verfahren (Expertensysteme). Die numerischen Methoden erlauben einem mehr oder weniger raffinierte Modelle online zu entwickeln, während die Expertensysteme als Interpretationshilfestellungen gedacht sind.

2.1 Kern-Dienstleistungen

2.1.1 Numerische Analysen

Alle bereits bekannten Analysemethoden werden ihren Platz unter den benutzerfreundlichen Modulen finden. Hierbei werden nur die eigenen, know how geschützten Lösungen kurz dargestellt:

Im Zentrum aller Methoden steht die Ähnlichkeitsanalyse. Es wird ständig nach potentiellen Mustern unter den Zahlen der bereitgestellten Objekt-Attribut-Matrizen gesucht. Die hierfür entwickelten Dienste unterscheiden sich hauptsächlich dadurch, wie viele Freiräume für das benutzergesteuerte Experimentieren zulassen.

2.1.2 Expertensysteme

Sobald die numerischen Analyseergebnisse vorliegen, kann der Anwender die Expertensysteme zur Hilfe rufen, welche nach passenden Schablonentexten suchen. Diese Schablonen sind Extrakte aus früheren Beratungsszenen und werden daher ständig weiter verfeinert und ergänzt.

Die Fragestellungen können völlig anonym, oder aber auch als gemeinnützig durch die Benutzer eingestuft werden. Die gemeinnützigen Fragestellungen sollten alle Objekt- und Attributbezeichnungen, sowie die eigenen Interpretationstexte (Berichte) detailliert freigeben, damit die Interpretationsexperten ständig nach neuen kontextunabhängigen Textmustern suchen können. Gemeinnützige Projekte helfen der Anwendergemeinschaft evt. Fehlinterpretationen zu filtern und gewisse kontextabhängige Ideen zu diskutieren.

2.2 Anwendungsbereiche

Die nachfolgenden Beispiele zeigen uns eindeutig, welche Bereiche der quasi online Entscheidungsunterstützung bereits existieren, jedoch nicht unbedingt optimal betrieben werden:

Sicherlich kennen viele bereits solche online Dienste wie z.B. www.preisvergleich.de. Hierbei kann man z.B. nach Kühlschränken anhand von Preisintervallen, Typen, Farben Herstellern, Maßen, etc. suchen. Die Filterungen sind zwar vielseitig, jedoch ersetzen nicht ganz die Autofilter-Potentiale einer Excel-Tabelle (lauter Kühlschränke und deren technischen und monetären Attribute). Das Endergebnis ist daher nicht anderes als eine Liste, wo man eine im Voraus ausgewählte Konstruktion am preiswertesten anschaffen kann. Das ist scheinbar eine komplette Lösung. Oder...?! Hätte man die Möglichkeit zu filtern (wie vorab beschrieben), wären sicherlich mehrere! (vergleichbare) Objekte, als potentiell kaufbaren Objekte übrig geblieben. Welche sichert jedoch das beste Preis-Leistungsverhältnis? Diese Antwort (unter anderem) liefert my_X auf Knopfdruck, ohne unnötige Fragen nach Gewichten und Punktezahlen je Attribut und Wert zu stellen... Handelt es sich nicht um Waren sondern Dienstleistungen, ist die Lösung ebenfalls gegeben. Hierbei denken Sie bitte immer daran: Man hat häufig in solchen Fällen zu entscheiden, wo die Angebote der potentiellen Lieferanten keine identischen, jedoch an sich passenden Leistungen beinhalten...

Viele Firmen stellen sich regelmäßig die Frage: Wie könnten die Kerngeschäfte besser gemacht werden. Benchmarking (z.B. <http://www.emendio.de>) liegt also immer im Trend. Falls Daten über mehrere Objekte (z.B. Betriebe, Verfahren) vorliegen (vgl. information brokering), dann kann man bereits mit Hilfe der my_X-Dienste danach suchen, welche Attribute für welche Erfolgskriterien wie weit verantwortlich sind.

Wird man freiwilliges Mitglied an den <http://www.meingehalt.com/> Diensten, so kann man erfahren, gibt es etwa solche Arbeitnehmer (wo, wie viele), die ähnliche! Kompetenzen/Aufgaben nachweisen und wie viel sie dabei verdienen. Sicherlich ist es sowohl für die Arbeitgeber als auch für die Arbeitnehmer zu wissen: Ist jemand lieber über-, oder unterbezahlt? Diese Antwort zu finden, braucht man eine multidimensionale Ähnlichkeitsanalyse (vgl. Preis-Leistungsvergleich oder Benchmarking).

Die neuesten TV-Sender in Ungarn haben (sehr richtig) solche Expertengespräche (endlich) im Angebot, wo darüber nachgedacht wird, (gesellschaftlich, wirtschaftlich, etc.) was erwartet einen z.B. in einer Woche? Diese Sendungen sind äußerst interessant, jedoch ohne echte Chancen, Trefferquoten für die Experten zu bilden (weder für die Vergangenheit, noch für die Zukunft). Wäre es nicht ideal, potentielle künstliche (Haus)-Experten zu „züchten“, dessen exakte Aussagen sofort geprüft und in Zukunft auch unter die Lupe genommen werden können? Um konkreter zu werden: Wie ändert sich die Inflation, die Arbeitslosigkeit, der Währungskurs, der Dieselpreis, die Temperatur, etc. wenn uns die Zeitreihen von vielen Erklärungsvariablen vorliegen? Warum sind lebendige Experten (die Nachfolger der „Zukunftswerkstatt von Dodona“) und deren vage Interpretationen glaubwürdiger, als solche, die durch die „Zahlenmystik“ direkt für uns nach unseren eigenen Qualitätsmerkmalen erstellt werden...

2.3 Wissenschaftliche Aspekte

Um die Schatten der bisherigen „Marketing-Dominanz“ teilweise zu kompensieren, sollte hier eine kurze Liste ernster, jedoch ziemlich vernachlässigter wissenschaftlichen Probleme stehen, welche sich zwar anscheinend nie lösen lassen, jedoch durch geschicktes Balancieren immerhin zu nützlichen Werkzeugen führen können:

Welche (Prognose)-Modelle sind besser in den nächsten wahren Entscheidungen?

Wann führen erhöhte Datenmengen zu besseren Modellen?

Welche abgeleiteten Zustandkombinationen der Zukunft sind als konsistent einzustufen?

Für welche Zeitintervalle kann man, für welche darf man nicht Prognose zu liefern?

Welche Muster sind am meisten kausal ohne experimentieren zu dürfen oder was muss man als Scheinkorrelation interpretieren?

Wie lohnt es eine gegebene Datenmenge zwischen Test- und Lerndaten zu verteilen?

Wie kann man sich gegen zu guten Modellen schützen?

Wie kann man am schnellsten beliebige Modelle in einem unendlichen Raum finden?

Welche Modellgütekriterien sind sparsam und robust genug, was die Rechenzeit und Wissenstransferstabilität betreffen? ...

3 Zusammenfassung

Ab Januar 2007 finden Sie also die erste Version der my_X Dienstleistung unter <http://www.my-x.hu>. Die geplanten Module erlauben den Anwendern anhand von eigentlich beliebigen Datenmengen Expertisen „real time & mobile“ ableiten zu lassen, ohne spezielle, eigene Ressourcen and Kompetenzen ausbauen zu müssen. Die Hintergründe der Berechnungen lassen sich durch e-learning tools visualisieren, so dass der Kunde niemals vor einem black box steht. Die angebotenen Annäherungen vertreten eine Art wirtschaftliche Optimallösung, welche in der gegebenen Zeit und durch die unternommenen Initiativen an der Seite des Anwenders von einer Maschine (welche extrahierte Gedanken vieler Experten konserviert - vgl. Schachautomat) zu erwarten ist.

Literaturverzeichnis

- [Pi04] Pitlik, L.: Component-based Object Comparison for Objectivity, 25. GIL - Jahrestagung, Bonn, 08-10. 09.2004, <http://miau.gau.hu/miau/69/gilfull.doc>
- [Pi05] Pitlik, L.; Bunkóczi, L.; Pető, I.: Environmental-Ecological Consistencies in automation of modelling (in: VII. Hungarian Biometric and Biomathematics Conference, Budapest 2005.07.05-06.), http://miau.gau.hu/miau/82/kjm_en_ecology_full.doc