

## Autoreferat

1. Imię i nazwisko: **Krzysztof Grąbczewski**

2. Posiadane dyplomy, stopnie naukowe

**magisterium** – 1994, Uniwersytet Mikołaja Kopernika w Toruniu, kierunek matematyka, specjalność informatyka,

**doktorat nauk technicznych** – 2003, Instytut Badań Systemowych PAN, Warszawa, rozprawa pt. "Zastosowanie kryterium separowalności do generowania reguł klasyfikacji na podstawie baz danych".

3. Informacje o dotychczasowym zatrudnieniu w jednostkach naukowych.

od	do	miejsce pracy	stanowisko
13.02.1995	30.09.2003	Uniwersytet Mikołaja Kopernika, ul. Gagarina 11, Toruń	asystent
1.10.2003	teraz	Uniwersytet Mikołaja Kopernika, ul. Gagarina 11, Toruń	adiunkt

4. Wskazanie osiągnięcia wynikającego z art. 16 ust. 2 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz. U. nr 65, poz. 595 ze zm.):

(a) tytuł osiągnięcia naukowego: **Meta-uczenie w indukcji drzew decyzji**

(b) autor, tytuł publikacji, rok wydania, nazwa wydawnictwa, **Krzysztof Grąbczewski, Meta-Learning in Decision Tree Induction, 2013, Springer.**

(c) omówienie celu naukowego ww. pracy i osiągniętych wyników wraz z omówieniem ich ewentualnego wykorzystania.

Głównym celem moich prac podejmowanych w ostatnich latach było opracowanie metod meta-uczenia w inteligencji obliczeniowej, a w szczególności w dziedzinie indukcji drzew decyzji. Uważam, że problem meta-uczenia jest aktualnie jednym z najważniejszych w dziedzinie i że jego atrakcyjne rozwiązania mogą w przyszłości pozwolić na bardzo istotny krok naprzód w kierunku autonomicznych systemów inteligentnych. Naukowcy zajmujący się dziedziną opracowali już mnóstwo algorytmów uczenia o różnych architekturach i własnościach, pozwalających na znajdowanie atrakcyjnych modeli dla różnych rodzajów danych. Obecnie, głównym problemem jest nie kwestia braku adekwatnych metod i potrzeba ich opracowania, a odnalezienie jak najskuteczniejszych metod uczących w gąszczu dostępnych powszechnie algorytmów. Dobór odpowiednich metod, które użyte we właściwej kombinacji i z właściwymi parametrami,

zaowocują dokładnymi, atrakcyjnymi modelami, jest zadaniem bardzo trudnym nawet dla ekspertów.

Podstawowe zadania meta-uczenia to automatyczne wyszukiwanie odpowiednich algorytmów uczących, gromadzenie meta-wiedzy o zaletach i wadach różnych metod, o okolicznościach, w których odpowiednie kombinacje metod prowadzą do dokładnych modeli oraz wykorzystywanie tej wiedzy w dalszych procesach uczenia. Każde z tych zadań może być rozwiązywane na przeróżne sposoby. Ani cel meta-uczenia, ani jego metodyka nie są określone wąsko i precyzyjnie, by mogły być łatwo i jednoznacznie realizowane. Nawet po bardzo istotnym zawężeniu zakresu badań (np. do określenia optymalnych postaci danych uczących dla konkretnej metody uczenia maszynowego, czy też do analizy skuteczności różnych konfiguracji parametrów danego algorytmu uczącego się), zadanie zwykle pozostaje NP-trudne, jednakże już dzisiaj znajdowanie atrakcyjnych, bliskich optymalnym, rozwiązań wydaje się być w zasięgu naszych możliwości (merytorycznie i obliczeniowo).

Z racji wcześniejszego zainteresowania algorytmami indukcji drzew decyzji, oraz ze względu na ich dużą popularność, uznałem za właściwe podjęcie tematu meta-uczenia szczególnie w odniesieniu do tego typu algorytmów.

Aby sprawnie podejmować kwestie meta-uczenia w dziedzinie drzew decyzji, potrzebne były bardzo duże nakłady pracy, by przygotować drogę do ostatecznego celu. Zabrały one sporo wysiłku i czasu, ale były niezbędne i otworzyły dużo nowych możliwości przyszłych badań. Wysiłki te można podzielić na kilka etapów:

- analiza istniejących algorytmów indukcji drzew decyzji pod kątem możliwości ich unifikacji,
- opracowanie ogólnego modelu algorytmów drzew decyzji pozwalającego na sprawne konstruowanie algorytmów przez kombinację kompatybilnych metod szczegółowych,
- rozwój ogólnych metod meta-uczenia tzn. mechanizmów jednolitego zarządzania parametrami algorytmów uczących się oraz dostępu do uzyskiwanych przez nie wyników,
- analiza eksperymentów porównawczych wykonanych z użyciem alternatywnych metod rozwiązywania poszczególnych podzadań,
- opracowanie algorytmów meta-uczenia, które w sposób w pełni automatyczny bądź półautomatyczny ekstrahują meta-wiedzę z wykonywanych eksperymentów obliczeniowych oraz wykorzystują tę meta-wiedzę w procesach doboru algorytmów uczących się dla konkretnych zadań.

Wszystkie wyliczone powyżej cele opisałem szerzej w książce pt. "Meta-Learning in Decision Tree Induction", którą przedkładałem jako rozprawę habilitacyjną. Rozprawa została przygotowana w taki sposób, by przedstawiać istotne elementy złożonej drogi do zaawansowanych algorytmów meta-uczenia. Będąc opisem prowadzonych przeze mnie badań, jest jednocześnie pomyślana jako spójne źródło wiedzy nt. metod indukcji drzew decyzji, którego wciąż brakowało na rynku publikacji mimo, że dostępne są tysiące artykułów poruszających różne drobne aspekty tej dziedziny.

Wiele popularnych i mniej znanych technik stosowanych w indukcji drzew decyzji zostało przeanalizowanych w rozdziale 2. Analiza ta stała się podstawą opracowania ogólnego schematu algorytmów, który pozwala wyodrębnić elementy konstrukcyjne algorytmów, pogrupować je według celów i stosowanych technik oraz składać te elementy

na mnóstwo sposobów w celu uzyskania różnych algorytmów budujących modele w postaci drzew decyzji. Tak uporządkowany przegląd oraz wyniki ogólny algorytm są przedstawione w trzecim rozdziale książki.

Sprawne posługiwanie się algorytmami indukcji drzew decyzji jest możliwe tylko wtedy, gdy oprócz metod konstrukcji drzew mamy również do dyspozycji narzędzia pozwalające na wykonywanie wielokrotnych testów, przygotowywanie danych do uczenia na wiele sposobów (różne transformacje danych) oraz estymację jakości powstałych modeli (np. szacowanie dokładności klasyfikatorów, wykonywanie testów statystycznych dla oceny istotności różnic itp.). Dlatego też, jednym z istotnych (i bardzo czasochłonnym) kroków w kierunku meta-uczenia było opracowanie ogólnego systemu do konstrukcji i testowania maszyn uczących się. W latach 1998-2005 współpracowałem z firmą FQS, dla której w naszym zakładzie stworzyliśmy system do analizy danych GhostMiner [43]. System ten jednak powstał pod wpływem innych przesłanek niż chęć meta-uczenia, więc nie był dość elastyczny u podstaw, by mógł służyć jako środowisko do szeroko zakrojonych badań nad meta-uczeniem. W związku z tym, pod wpływem wspólnych zainteresowań meta-uczeniem, wraz z kolegą, Norbertem Jankowskim, zaprojektowaliśmy i zaimplementowaliśmy nowy system o nazwie Intemi, który uwzględniał potrzeby meta-uczenia od początkowej fazy projektu. Pozwala on w jednolity sposób operować różnymi algorytmami uczącymi się, przetwarzającymi dane, wykonującymi testy itp., bez konieczności posiadania wiedzy na temat wewnętrznej struktury tych algorytmów. Założenia architektury systemu Intemi oraz jego najbardziej potrzebne dla meta-uczenia elementy są przedstawione w rozdziale 4 mojej książki.

System Intemi stał się bazą dla implementacji uogólnionego modelu drzew decyzji i licznych testów przeprowadzonych na użytek meta-uczenia. Ogólny model został wyposażony w szereg komponentów, które pozwalają realizować mnóstwo algorytmów indukcji drzew decyzji (tych opisanych w rozdziale 2 książki oraz wielu innych). W ten sposób powstał najbogatszy z dotychczas opublikowanych systemów do indukcji drzew decyzji (pozwalający realizować największą liczbę powszechnie znanych algorytmów, a także łatwo składać nowe algorytmy z gotowych komponentów). Wybrane wyniki testów porównujących różne składowe algorytmów zostały przedstawione w rozdziale 5 książki. Można tam znaleźć szereg istotnych wniosków nt. przydatności poszczególnych komponentów dla indukcji atrakcyjnych drzew decyzji.

Najważniejsze z moich dotychczasowych wysiłków w zakresie meta-uczenia znalazły opisy w rozdziale 6. Polegają one na kontrolowanym szukaniu w przestrzeni (parametrów) algorytmów. Książka przedstawia dwa podejścia zaprojektowane dla zastosowań w różnych przypadkach: pierwsze oparte jest na testowaniu maszyn uczących w odpowiednim porządku uwzględniającym m.in. złożoności maszyn, drugie jest przeznaczone dla poszukiwania wśród algorytmów nie różniących się istotnie złożonością—tutaj analiza profili wyników uczenia decyduje o kolejności przeprowadzania testów. Algorytmy te zostały zdefiniowane w sposób ogólny, pozwalający z dużą elastycznością definiować różne rozwiązania szczegółowe, poprzez podstawianie odpowiednich komponentów. Dzięki eksploracji przestrzeni maszyn uczących, takie metody pozwalają gromadzić meta-wiedzę niezbędną dla poprawnego rozwiązywania problemów doboru metod uczących się do analizowanych danych. W ostatnich latach coraz popularniejsze staje się tworzenie tzw. ontologii, czyli specjalnych repozytoriów wiedzy, na użytek meta-uczenia. W przyszłości powinno to pozwolić na transfer wiedzy pozyskanej w procesach meta-uczenia na nowe zadania, co zaowocuje skróceniem czasu poszukiwania atrakcyjnych rozwiązań lub, patrząc z innej perspektywy, pozwoli znajdować dokład-

niejsze modele w określonym czasie, co jest celem nadrzędnym wszelkich procesów meta-uczenia.

Jak dotąd, wciąż nie dysponujemy w pełni automatycznymi metodami wyszukującymi modele klasyfikacji, aproksymacji czy grupowania danych, które dorównywałyby osiągnięciami ludzkim ekspertom. Dalszy rozwój meta-uczenia, liczne eksperymenty przy użyciu takich algorytmów jak te prezentowane w rozdziale 6 książki oraz budowanie odpowiednich ontologii i mechanizmów sprawnego dostępu do zawartej w nich wiedzy powinny pozwolić w przyszłości na uzyskanie algorytmów uczących się przewyższających możliwościami eksperta-człowieka.

Kolejne generacje algorytmów meta-uczenia powinny umożliwić tworzenie systemów uczących się z danych, których użytkownik będzie musiał jedynie udostępnić dane wejściowe i ewentualnie określić, czym dokładniej jest zainteresowany (np. maksymalną dokładnością klasyfikacji, maksymalną wrażliwością, zrozumiałym opisem decyzji w postaci reguł logicznych), po czym po prostu uruchomi system i będzie na bieżąco informowany o znalezionych przez system modelach. Uważam, że już obecnie, moc obliczeniowa i dostępne w komputerach ilości pamięci (operacyjnej i masowej) pozwalają na tworzenie takich systemów, a głównym powodem ich braku jest brak całościowych implementacji maszyn meta-uczących się i wykorzystujących odpowiednio zorganizowane ontologie. Środowisko Intemi, opisane w książce, zostało stworzone właśnie po to, aby umożliwić konstruowanie takich złożonych systemów. Dotychczasowe implementacje algorytmów meta-uczenia potwierdzają jego możliwości i pozwalają z optymizmem oczekiwać wspomnianych wyżej rozwiązań.

## 5. Omówienie pozostałych osiągnięć naukowo-badawczych.

### **Przed uzyskaniem stopnia doktora:**

#### ***Automatyczne dowodzenie twierdzeń***

W ramach pracy magisterskiej oraz na początku pracy naukowej zajmowałem się zagadnieniami automatycznego dowodzenia twierdzeń. Współpracowałem wówczas z drem Lawrence C. Paulsonem z Uniwersytetu w Cambridge, gdzie przebywałem dwukrotnie na stażach w 1994 i 1995 roku (odpowiednio 7 miesięcy i 3 miesiące). Moja praca dotyczyła głównie problemów maszynowego dowodzenia równoważności różnych sformułowań aksjomatu wyboru oraz własności dobrego uporządkowania [1, 2].

#### ***Ekstrakcja reguł logicznych z danych***

Istotną część mojej aktywności naukowej przed uzyskaniem stopnia doktora były badania nad metodami ekstrakcji reguł logicznych z danych. Moje publikacje w tym temacie [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 16, 17, 18, 19, 20, 21, 23, 24, 25, 26, 27, 28, 29, 30, 31] dotyczyły metod neuronowych (specjalnych architektur sieci pozwalających na konwersję sieci do postaci interpretowalnej jako reguły logiczne), metod szukania odpowiednich struktur sieci i parametrów węzłów, metod optymalizacji reguł logicznych, a także metod hybrydowych łączących różne podejścia.

#### ***Algorytmy indukcji drzew decyzji***

Zajmowałem się również algorytmami indukcji drzew decyzji, w ramach których opracowałem kryterium separowalności na użytek tworzenia drzew oraz różne dodatkowe metody niezbędne do indukcji i walidacji pojedynczych drzew decyzji i kolekcji drzew (lasów), oraz do budowania heterogenicznych drzew łączących w sobie decyzje oparte na klasycznych podziałach i decyzje minimalnoodległościowe

[15, 22, 32, 36, 37]. Metody te stały się tematem mojej rozprawy doktorskiej [39] obronionej w 2003 roku w Instytucie Badań Systemowych PAN.

### ***Selekcja cech***

Pracowałem także nad metodami selekcji istotnych cech opisujących dane, w celu zapewnienia lepszych wyników działania różnych klasyfikatorów. Badania obejmowały przede wszystkim metody oparte na teorii informacji oraz odpowiednio zaadaptowane metody podziałów węzłów drzew decyzji [34, 35].

### ***System GhostMiner***

Algorytmy opracowane przez kilka osób z naszego zespołu zostały zorganizowane w jednolity system GhostMiner ([www.fqs.pl/ghostminer](http://www.fqs.pl/ghostminer)), który rozwijaliśmy we współpracy z firmą FQS z Krakowa (firmą-córką japońskiej Fujitsu) [38]. System ten znalazł wielu odbiorców na całym świecie wśród instytucji naukowych i komercyjnych. Na jego bazie, firma FQS nadal tworzy liczne aplikacje biznesowe. Na użytek systemu GhostMiner, opracowaliśmy również system dokumentacji i pomocy kontekstowej do jednorodnego tworzenia obu rodzajów dokumentów na podstawie wspólnego źródła [33].

### **Po uzyskaniu stopnia doktora:**

Moje zainteresowania naukowe po uzyskaniu stopnia doktora przesunęły się w kierunku algorytmów meta-uczenia maszynowego, czyli metod wyboru najbardziej odpowiednich maszyn uczących oraz prób zautomatyzowania wyciągania wniosków nt. użyteczności poszczególnych metod w konkretnych zastosowaniach. W związku z tym, pracowałem nad szerokim spektrum algorytmów uczących się z danych oraz metod transformacji danych, jako bardzo istotnego elementu tworzenia docelowych modeli. Jednocześnie nadal interesowałem się metodami indukcji drzew decyzji, bardzo wartościowymi zarówno w uczeniu maszynowym na poziomie bazowym (klasyfikacji, aproksymacji, klasteryzacji itd.) jak i na poziomie meta (analizy działania algorytmów uczących się z danych). Potrzeby meta-uczenia wymusiły stworzenie nowego systemu (o odpowiednio opracowanej architekturze) do tworzenia maszyn uczących i analizy ich procesów uczenia i indukowanych przez nie modeli. Stworzenie tego systemu pozwoliło opracować i analizować różne algorytmy meta-uczenia.

### ***Metody uczenia maszynowego na poziomie bazowym***

Ponieważ sprawne uczenie na poziomie meta wymaga możliwości tworzenia modeli różnych typów na poziomie bazowego uczenia się oraz możliwości sprawnego zarządzania parametrami uczenia maszyn na poziomie bazowym, analizowałem i rozwijałem metody uczenia oparte na różnych paradygmatach, np. Support Vector Machines, metody minimalnoodległościowe, Bayesowskie i mieszane [41, 48, 42, 44, 46]. Metody te obejmowały również komitety z oceną kompetencji poszczególnych członków, z uwzględnieniem oceny jakości ich decyzji dla całej przestrzeni danych oraz w jej podobszarach. Oprócz kompetencji (lokalnej i globalnej) analizie poddawałem również efektywności różnych metod dla konkretnych danych i różne wysiłki zrozumienia danych. Dzięki podejmowaniu analizy szerokiego spektrum metod uczących się z danych, możliwe było przygotowanie zwięzłego, ale bogatego w informacje przeglądu cech i możliwości różnych algorytmów uczenia maszynowego, który został opublikowany jako pierwszy rozdział książki nt. metod ekstrakcji cech [47] opublikowanej po konkursie zorganizowanym w ramach konferencji NIPS 2003 (p. niżej).

### **Transformacje danych**

Od początku prowadzenia badań naukowych w dziedzinie uczenia maszynowego, opowiadałem się przeciwko stosowaniu wyraźnego podziału procesu analizy danych na wstępne przetwarzanie danych oraz finalne uczenie modelu decyzyjnego. Transformacje danych są często bardziej istotne dla ostatecznego sukcesu niż algorytmy budujące modele w ostatnim etapie uczenia, dlatego są one szczególnie istotne z punktu widzenia meta-uczenia i muszą być traktowane jako integralna część maszyny uczącej się z danych. Wydzielanie ich jako etap wstępnego przygotowania danych prowadzi często do przekłamań wyników i wniosków wyciąganych z eksperymentów.

Podjewałem szereg prac poddających analizie metody transformacji danych i tworzących nowe transformacje, włącznie z algorytmami ekstrakcji cech, a w szczególności selekcji informatywnych cech, tworzenia nowych cech poprzez odpowiednie transformacje oryginalnych danych, konwersji cech dyskretnych na ciągłe (na użytek modeli decyzyjnych wymagających ciągłości cech), konwersji cech ciągłych na dyskretne itp. [40, 42, 45, 48, 53]. Wartość prac nad transformacjami danych potwierdziły wyniki konkursów analizy danych, w których brałem udział (p. niżej).

### **Konkursy analizy danych**

Analiza różnych metod uczenia maszynowego, ze szczególnym uwzględnieniem metod transformacji danych, pozwoliła zakończyć sukcesem udział w konkursach analizy danych poświęconych selekcji cech na użytek klasyfikacji [48] oraz rozpoznawania pisma ręcznego [53]. W pierwszym z tych konkursów, zorganizowanym w ramach konferencji NIPS 2003, zajęliśmy (przy współudziale Norberta Jankowskiego) trzecie miejsce, a drugi, zorganizowany przy konferencji ICAISC 2006, wygraliśmy. W obu bardzo istotne okazało się opracowanie odpowiednich metod przetwarzania danych przed budową ostatecznych modeli decyzyjnych. Sukces w konkursie NIPS 2003 zaowocował zaproszeniem do współautorstwa książki poświęconej zagadnieniu ekstrakcji cech, w której znalazły się dwa nasze rozdziały [47, 48], pierwszy ogólny na temat uczenia maszynowego, a drugi o algorytmach opracowanych na użytek konkursu.

Wydaje mi się ważnym podkreślenie, że od uzyskania stopnia doktora, wzięłem udział tylko w tych dwóch wspomnianych wyżej konkursach, więc wspomniane wyniki nie są najlepszymi wybranymi z wielu prób, a dają kompletny obraz porównawczy opracowanych przeze mnie metod oraz tych stosowanych przez innych badaczy, którzy też wzięli udział w tych konkursach.

### **Indukcja drzew decyzji**

Z racji wcześniejszych zainteresowań algorytmami indukcji drzew decyzji i dobrej ich znajomości, poddawałem takie algorytmy dalszym badaniom, zwłaszcza w kontekście ich przydatności dla meta-uczenia maszynowego [45, 61, 62, 63, 66]. Badania te stały się bardzo istotną częścią badań opisanych w monografii niniejszym przedkładanej jako główne osiągnięcie dla postępowania habilitacyjnego.

### **Systemy GhostMiner oraz Intemi**

Osiągnięcie celów meta-uczenia wymagało posiadania niezwykle elastycznego i efektywnego systemu do wszechstronnego uczenia maszynowego. Początkowo rozwijaliśmy jeszcze system GhostMiner [43], jednakże potrzeby meta-uczenia wymagały stosowania innej, ogólniejszej architektury u podstaw systemu. Efektem tego była decyzja o zaprojektowaniu od początku i zbudowaniu nowego systemu, który

otrzymał nazwę *Intemi* [50, 52, 54, 58, 59, 60]. System ten wydaje się być obecnie najbardziej wszechstronnym dostępnym narzędziem do uczenia maszynowego, choć niestety, głównie z powodu braku czasu na przygotowanie odpowiednio wygodnego interfejsu użytkownika i wyczerpującej dokumentacji, nie jest on obecnie udostępniany innym badaczom niż bliscy współpracownicy naszego zespołu.

### **Meta-uczenie maszynowe**

Mając do dyspozycji system *Intemi*, można realizować bardzo złożone projekty uczenia maszynowego, w szczególności te podejmujące meta-uczenie. W ostatnim czasie, pracowałem nad kilkoma zaawansowanymi metodami meta-uczenia [49, 51, 55, 56, 57, 65, 67], z których dwie podstawowe (a właściwie dwa ogólne schematy metod, którymi można realizować różne cele) opisałem w przedkładanej monografii [68].

Naszemu zespołowi udało się także zgromadzić badaczy z całego świata, zajmujących się różnymi aspektami meta-uczenia wokół wspólnego celu przygotowania książki przedstawiającej aktualny stan badań w tej dziedzinie [64].

*Ryszard Gubian*