

ПРОГРАМА
вступного іспиту зі спеціальності
186 «Видавництво та поліграфія»
для вступників на навчання в аспірантурі

На вступні випробування виносяться блоки таких фахових дисциплін:

- «Інтелектуальні методи обробки цифрового контенту видавничих систем»
- «Проблеми управління великими даними в задачах аналізу медіаконтенту»;
- «Аналіз неструктурованих даних у Веб-системах »;
- «Методи аналізу та обробки текстів»;

Інтелектуальні методи обробки цифрового контенту видавничих систем

Основні поняття машинного навчання. Приклади прикладних задач, які використовують методи машинного навчання. Методи класифікації, регресійного аналізу, кластеризації, пошуку в просторі гіпотез. Задачі прогнозування та прийняття рішень. Задачі оптимізації — генетично-адаптивні алгоритми. Метод опорних векторів, метод «найближчого сусіда» і байєсівський методи класифікації. Дерева рішень. Нейромережеві методи класифікації та регресії — навчання на основі зв'язків. Нейромережеві методи класифікації та регресії — мережі, що самоорганізуються. Байєсові мережі. Пошук асоціативних правил — навчання без вчителя. Інтелектуальні агенти та мультиагентні системи.

Задачі Data Mining. Застосування технології Data Mining. Поняття Web Mining, Text Mining, Call Mining, Opinion Mining. Способи візуального представлення даних. Методи візуалізації. Етапи процесу Data Mining. Інструменти Data Mining.

Рекомендована література:

1. Д.Мейер Теория реляционных баз данных: Пер. с англ. — М.: Мир, 1987. — 608 с., ил.
2. Берко А.Ю., Верес О.М., Пасічник В.В. Системи баз даних та знань. Книга 1. Організація баз даних та знань: Навчальний посібник. — Львів: «Магнолія 2006», 2008. — 456 с.
3. К. Дж. Дейт. Введение в системы баз данных, 8-е изд.: Пер. с англ. — М.: Издательский дом «Вильямс», 2005. — 1328 с.: ил. — Парал. тит. англ.
4. Пасічник В. В. Організація баз даних та знань. / В. В. Пасічник, В. А. Резніченко. — К.: Видавнича група ВНУ, 2006. — 384 с.
5. Ульман Дж. Основы систем баз данных. Пер. с англ. — М.: Финансы и статистика, 1983. — 334 с.

Аналіз неструктурованих даних у Веб-системах

Поняття контент аналізу, категорії аналізу, методи контент аналізу, методики відбору джерел аналізу, емпіричні моделі аналізу, методики інтерпретації даних. Кількісний та якісний контент аналіз. Процедури оптимізації зберігання. Виявлення непотрібних даних і

позбавлення від них при міграції. Класифікація. Дотримання нормативів і вимог (compliance). Керування рівнями доступу. Text Mining для аналізу неструктурованих даних. Автоматизація видобування інформації з неструктурованих даних. Когнітивний аналіз неструктурованих даних. Ієрархічні методи аналізу. Поняття Web Mining. Інструменти Web Mining/

Рекомендована література:

1. Барсегян А.А., Куприянов М.С. и др. Анализ данных и процессов — 3-е изд., перераб и доп. — СПб.: БХВ-Петербург, 2009. — 512 с.: ил. + CD-ROM.
2. Иванов В. Ф. Контент-анализ. — К., 1994. — 110 с..
3. Костенко Н, Иванов В. Досвід контент-аналізу: моделі та практики: Монографія. — К.: Центр вільної преси, 2003. — 44 с..
4. Thomas, R. W., Friend, D. H., DaSilva, L. A., MacKenzie, A. B. (2006). Cognitive Networks: Adaptation and Learning to Achieve End-to-end Performance Objectives. IEEE Communications Magazine, 12 (44), 21.
5. Checkland, P. B. (1981). Systems Thinking, Systems Practice. New York: Wiley, 630.
6. Avdeeva, Z. K., Kovriga, S. V., Makarenko, D. I. (2010). Cognitive modeling for solving semi-structured management systems (situations). Institute of Control Sciences, 26–39

Проблеми управління великими даними в задачах аналізу медіаконтенту

Поняття великих даних. методи класу Data Mining для управління великими даними - навчання асоціативним правилам (association rule learning), класифікація (методи категоризації нових даних на основі принципів, раніше застосованих до вже наявним даними), кластерний аналіз, регресійний аналіз; Краудсорсинг - категоризація даних; Інтеграція даних (data fusion and integration); Використання машинного навчання, використання моделей, побудованих на базі статистичного аналізу або машинного навчання для отримання комплексних прогнозів на основі базових моделей (constituent models), зі статистичним ансамблем у статистичній механіці; штучні нейронні мережі, мережевий аналіз, оптимізація, в тому числі генетичні алгоритми; розпізнавання образів; прогнозна аналітика; імітаційне моделювання; просторовий аналіз (Spatial analysis); візуалізація аналітичних даних.

Рекомендована література:

1. Барсегян А.А., Куприянов М.С. и др. Анализ данных и процессов — 3-е изд., перераб и доп. — СПб.: БХВ-Петербург, 2009. — 512 с.: ил. + CD-ROM.
2. CRISP-DM 1.0. Step-by-step data mining guide. — SPSS, 2000.
3. Artificial Intelligence — A Guide to Intelligent Systems, Michael Negnivitsky, Addison-wesley, Pearson Education Limited? 2002.
4. Люгер Д. Искусственный интеллект. — М.: Мир, 2003. — 690 с.
5. Айвазян С. А., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: основы моделирования и первичная обработка данных. — М.: Финансы и статистика, 1983.
6. Айвазян С. А., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: исследование зависимостей. — М.: Финансы и статистика, 1985.
7. Айвазян С. А., Бухштабер В. М., Енюков И. С., Мешалкин Л. Д. Прикладная статистика: классификация и снижение размерности. — М.: Финансы и статистика, 1989.
8. Вапник В. Н. Восстановление зависимостей по эмпирическим данным. — М.: Наука, 1979.

9. Журавлев Ю. И., Рязанов В. В., Сенько О. В. «Распознавание». Математические методы. Программная система. Практические применения. — М.: Фазис, 2006. ISBN 5-7036-0108-8.
10. Загоруйко Н. Г. Прикладные методы анализа данных и знаний. — Новосибирск: ИМ СО РАН, 1999. ISBN 5-86134-060-9.
11. Шлезингер М., Главач В. Десять лекций по статистическому и структурному распознаванию. — Киев: Наукова думка, 2004. ISBN 966-00-0341-2.
12. Николенко С.И. Курс лекций по машинному обучению — слайды. Электронный ресурс. Режим доступа: <http://logic.pdmi.ras.ru/~sergei/index.php?page=mlaptu09>

Методи аналізу та обробки текстів

Технологія аналізу текстів. Завдання Text Mining. Основні складові Text Mining - Information Extraction (видобування інформації, зокрема видобування ознак, зв'язків між ознаками, видобуток сутностей та аномалій); Summarization (автоматичне реферування, анотування); Categorization (категоризація, класифікація); Clusterization (кластеризація). Проблеми комп'ютерного аналізу та синтезу текстів природної мови (Natural Language processing). Головні завдання обробки текстів природної мови та методи їх розв'язання. Задачі та обмеження обробки текстів природної мови. Рівні обробки природної мови

Рекомендована література:

1. ЛАНДЭ Д. ГЛУБИННЫЙ АНАЛИЗ ТЕКСТОВ ТЕХНОЛОГИЯ ЭФФЕКТИВНОГО АНАЛИЗА ТЕКСТОВЫХ ДАННЫХ [Електронний ресурс] / Дмитрий ЛАНДЭ – Режим доступа до ресурсу: [http://visti.net/~dwl/art/dz/.](http://visti.net/~dwl/art/dz/)
2. Анисимов А. В., Марченко А. А. Система обработки текстов на естественном языке // «Штучний інтелект». — 2002. — № 4. -С. 157
3. Вгору ↑ Диковицкий В. В., Шишаев М. Г. Обработка текстов естественного языка в моделях поисковых систем // Сборник научных трудов. — 2010. — С.30.
4. Liddy, E.D. 2001. Natural Language Processing. In Encyclopedia of Library and Information Science, 2nd Ed. NY. Marcel Decker, Inc. — P.1
5. Вгору ↑ Иванов О. В. Класичний контент-аналіз та аналіз тексту: термінологічні та методологічні відмінності / Иванов Олег Валерійович // Вісник Харківського національного університету імені В. Н. Каразіна, Харків: Видавничий центр ХНУ імені В. Н. Каразіна, 2013. — № 1045. — С.72