

Программа спецкурса для аспирантов по специальности 05.13.17 «Теоретические основы информатики (концептуальные модели и математические основы)»

(полугодовой курс, 36 часов)

*Подготовили: д.ф.-м.н., проф. В.А.Васенин, к.ф.-м.н., вед.н.сотр. С.А.Афонин,
к.ф.-м.н., вед.н.сотр. А.С.Козицын,
к.ф.-м.н., вед.н.сотр. А.С.Шундеев, к.ф.-м.н., с.н.с. Д.Д.Голомазов*

Введение (2 часа)

Информатика как наука, изучающая информацию и ее свойства в естественных, искусственных и гибридных системах. Место информатики в системе наук. Информатика как обрабатывающая информацию отрасль индустрии и инфраструктурная область, ее роль и значение в ускорении научно-технического прогресса.

Понятие информационного продукта и информационной услуги (2 часа)

Классификация информационных продуктов и услуг. Жизненный цикл информационного продукта. Экономика информационных сетей. Методы управления производством и распределением информационных продуктов

Информационные ресурсы (2 часа)

Принципы оценки информации как ресурса общества и объекта интеллектуальной собственности. Проблемы правового регулирования научной интеллектуальной собственности. Государственная политика в области защиты информационных ресурсов общества.

Машинное обучение (4 часа)

Задача машинного обучения. Объекты и признаки. Основные понятия: метод обучения, функционал качества, обобщающая способность, скользящий контроль. Алгоритмы классификации: C4.5, анализ формальных понятий, метод опорных векторов, k ближайших соседей, Байесовские классификаторы, AdaBoost, скрытые модели Маркова, метод условных случайных полей. Факторный анализ. Алгоритмы кластеризации: k-средних, самоорганизующиеся карты Кохонена, графовые алгоритмы, иерархическая кластеризация.

Методы анализа текстовых данных (4 часа)

Приложения задач анализа текстовых данных: кластеризация, извлечение данных, выявление трендов. Алгоритмы выделения именованных сущностей, гиперонимов, описаний объектов. Разрешение неоднозначности. Алгоритмы выявления ассоциативных правил: a-priori, FP-growth.

Обработка слабоструктурированных данных (2 часа)

Графовые и древовидные модели данных. Понятие путевого запроса. Языки запросов XPath и XQuery. Методы выполнения запросов.

Алгоритмы анализа социальных сетей (2 часа)

Формализация понятия сообщества в социальной сети. Алгоритмы нахождения сообществ. Оценка авторитетности ресурсов. Алгоритм PageRank.

Приближенные алгоритмы дискретной оптимизации. (2 часа)

Примеры оптимизационных задач: минимальное трансверсальное множество, минимальное покрывающее множество. Жадные алгоритмы. Сведение к задаче линейного программирования. Алгоритмы со случайным выбором.

Методы представления знаний (2 часа)

Базы знаний. Общие принципы моделирования окружающей среды и мышления человека. Методы представления знаний: классификационные, тезаурусные, основанные на отношениях, семантические сети и фреймы, продукционные и непродукционные.

Онтологии (2 часа)

Введение в дескриптивную логику. Семейства логик. Характерные задачи. Логический вывод. Онтологии. Языки описания онтологий. Языки запросов к онтологиям.

Семантическая паутина (Semantic Web) (2 часа)

Введение в Семантическую паутину (Semantic Web). Краткая история развития Всемирной паутины. Основные технологии Semantic Web. Структура SW-приложения. Два подхода к реализации. Проект Linking Open Data. Перспективы развития Semantic Web.

Комплексного подхода к обеспечению информационной безопасности (2 часа)

Выделение уровней комплексного подхода к обеспечению ИБ: законодательный, административно-организационный, операционный и программно-технический. Обзор мер и методов на уровнях комплексного подхода к обеспечению ИБ.

Модели угроз и нарушителей (2 часа)

Основные принципы построения систем защиты информации в компьютерных системах. Модель угроз и модель нарушителя. Распространенные угрозы и атаки. Проектирование и выбор архитектуры системы защиты. Автоматизированные системы в защищенном исполнении.

Аутентификация (2 часа)

Идентификация и аутентификация в компьютерных системах. Общие положения создания систем и механизмов идентификации и аутентификации. Примеры реализации методов идентификации и аутентификации в операционных системах, в веб-приложениях.

Криптография и криптоанализ (2 часа)

Обзор решаемых задач и основных методов. Понятия криптографической системы, криптографического алгоритма, криптографического протокола. Обзор нормативной базы в области криптографии. Обзор типовых областей применения криптографических систем, алгоритмов и протоколов. Криптографическая стойкость. Обзор распространенных методов атак.

Критерии оценки информационных технологий и систем (2 часа)

Оценки качества поиска (полнота, точность и др.). Скалярные и векторные оценки. Смешанные критерии (полезная работа, корреляционный критерий, свертки и пр.). Рабочие характеристики информационно-поисковых систем (ИПС) в различных координатах. Вероятностная модель ИПС. Теоретико-множественная модель ИПС. Оптимизация режима ИПС. Линейное представление документов, запросов, индексирования, поиска.