

Аннотация рабочей программы дисциплины «Большие данные»

Цель дисциплины - изучение математических методов и моделей, используемых в системах обработки и анализа больших данных для поддержки принятия решений, и развитие профессиональных навыков в этой области

Задачи дисциплины:

- формирование теоретических и методологических основ в области анализа неструктурированной информации;
- формирование практических навыков, использования алгоритмов интеллектуального анализа данных.
- формирование навыков проведения сравнительного анализа основных моделей.

Названия тем, основных вопросов.

1. Группировка данных (понятие данных, совокупность инструментов, подходов и методов обработки больших данных)
2. Обнаружение значимых корреляций (понятие корреляции, средства массово-параллельной обработки, прогнозная аналитика)
3. Определение отношений между разнородными данными (распознавание объектов, имитационное моделирование, пространственный анализ)
4. Систематизированные структуры данных (Анализ временных рядов, коинтеграция, адаптивные и мультипликативные методы прогнозирования)
5. Вывод правил для принятия решений (факторный анализ, ранговые методы, неметрические методы, кластерный анализ)
6. Прогнозирование последствий принятых решений (прогнозная аналитика, нейронные сети, пространственный анализ, сетевой анализ)
7. Регрессионный анализ (понятие и способы применения регрессионного анализа, инструментарий для реализации)
8. Дисперсионный анализ (понятие и способы применения дисперсионного анализа, инструментарий для реализации)
9. Кластерный анализ (понятие и способы применения кластерного анализа, инструментарий для реализации)
10. Факторный анализ (понятие и способы применения факторного анализа, инструментарий для реализации)
11. Модели распределенных файловых систем (data mining, краудсорсинг, смешение и интеграция данных)
12. Поиск подобий в данных (машинное обучение, оптимизация и генетические алгоритмы)
13. Анализ потоковых данных (анализ связей, визуализация аналитических данных)
14. Социально-сетевые графы (частые наборы данных, прогнозирование и предвидение в социальных и медиа процессах)
15. Методы кластеризации (модели снижения размерности данных, методы машинного обучения большими данными)
16. Способы применения нейронных сетей (приложения нейронных сетей, сетевые аналитические модели)

Объем дисциплины 4 з.е.

Форма промежуточного контроля – экзамен.