Специальный курс по выбору студента

Кафедры математических и компьютерных методов анализа

**Большие данные в финансовых технологиях (1/2 г.)**

Лекторы: доцент О.В. Попов, приглашенные докладчики.

**Аннотация курса:** что такое Биг Дата, какие задачи она решает; почему сегодня эта задача актуальна в it-индустрии; Биг Дата в точки зрения разработчики по и с точки зрения специалиста; какие задачи в области Биг Дата решает программист, какие - специалист; техническая поддержка для решения задач Биг Дата: инструментальные средства, встраиваемые в цикл разработки по; демонстрация незащищенность программных приложений на реальных примерах.

**Программа курса**

Тема 1 Что такое большие данные? Каковы признаки больших данных?

Тема 2 Модель вычислений MapReduce. Устройство, назначение.

|  |
| --- |
| Тема 3 Распределённые файловые системы. HDFS. |
| Тема 4 HDFS. Примеры решаемых задач, текущее состояние. |
| Тема 5 NoSQL субд. Возникновение и назначение. Column-family на примере HBase. |
| Тема 6 SQL-подобные системы. Hive. Устройство и назначение. |
| Тема 7 SQL-подобные системы. Impala. Устройство и назначение. |
| Тема 8 Spark. Модель вычислений. Разница между Spark и MapReduce |
| Тема 9 Spark. Приёмы оптимизации. |
| Тема 10 Оптимизированные форматы хранения данных. |
| Тема 11 Поточная обработка данных. Kafka. Устройство платформы. |
| Тема 12 Kafka. Подходы к обработке данных. |
| Тема 13 Распределённые реляционные СУБД. Greenplum. |
| Тема 14 Архитектуры систем обработки данных - Lambda, Kappa. Традиционные DWH. |
| Тема 15 Архитектуры систем обработки данных - традиционные DWH. |

Форма промежуточной аттестации – экзамен.