Регламент курса «**АНАЛИЗ ИНТЕРНЕТ-ДАННЫХ**» для магистратуры «Прикладная математика и информатика», специализация «Анализ данных»

Курс «**АНАЛИЗ ИНТЕРНЕТ-ДАННЫХ**» читается в магистратуре по программе «Бизнес-информатика», 2 курс, 3 семестр.

Для освоения данного курса студенты должны изучить курсы по программе бакалавриата «Бизнес-информатика»:

* «Анализ данных»
* «Интернет-технологии»
* «Базы данных»,

а также иметь навыки работы в среде R.

Курс включает 16 ч лекций и 32 ч. лабораторных занятий.

На лекциях будут рассматриваться темы:

* Data mining и Web mining
* Основные Интернет-технологии для Web mining
* Технология R в Web Mining
* Text mining
* Социальные сети. Social mining
* Рекомендательные системы
* Интеллектуальные агенты
* Принципы функционирования поисковых систем

На лабораторных занятиях и в рамках самостоятельной работы студенты будут выполнять индивидуальные задания.

Балльно-рейтинговая система:

*Текущая работа в семестре:*

**Задание 1.** Анализ Web-логов средствами R (максимум **10 баллов**).

**Задание 2.** Загрузка данных из Интернета средствами специализированных программ и визуализация данных в R (максимум 10 баллов).

**Задание 3.** Загрузка данных из Интернета средствами R и анализ данных в R (максимум **20 баллов**).

**Дополнительное задание 1.** Использование баз данных в любом из 3 основных заданий (максимум 10 баллов).

*К зачету:*

**Задание 4.** Text mining средствами R (максимум **20 баллов**).

**Задание 5.** Разработка интеллектуального агента(максимум 30 баллов = **10 баллов за техническое задание** + 20 баллов за реализацию).

**Жирным** шрифтом выделены баллы за обязательные задания, на «троечку» **(30 баллов за текущую работу и 30 баллов за экзамен).**

Любой студент может пропустить в течение семестра **без всяких объяснений и справок** 5 занятий. Начиная с 6-го пропущенного занятия, начисляются штрафные баллы - по минус 2 балла за занятие.