

## **Практическое занятие 12. Классификация и кластеризация в текстовом анализе**

### **План практического занятия:**

- ✓ Классификация текста
- ✓ Перекрестная проверка
- ✓ Конструирование модели
- ✓ Обучение на текстовых данных без учителя
- ✓ Кластеризация документов по сходству
- ✓ Моделирование тематики документов
- ✓ Латентное размещение Дирихле
- ✓ Латентно-семантический анализ

### **Задания:**

1. Обзор научно - учебной информации по теме
2. Подготовьте конспект по плану практического занятия

### **Рекомендуемые источники информации:**

1. Бринк Х., Ричардс Д., Феверолф М. Машинное обучение. -СПб.: Питер, 2017. -336 с
  2. Автоматическая обработка текстов на естественном языке и анализ данных : учеб. пособие / Большакова Е.И., Воронцов К.В., Ефремова Н.Э., Клышинский Э.С., Лукашевич Н.В., Сапин А.С. — М.: Изд-во НИУ ВШЭ, 2017. — 269 с
  3. Коэлью, Л.П. Построение систем машинного обучения на языке Python // Л.П. Коэлью, В. Ричарт ; пер. с англ. Слинкин А. А. - Москва : ДМК Пресс, 2016. — 302 с.
- Воронина В В и др. Теория и практика машинного обучения: учебное пособие/Ульяновск: УлГТУ, 2017. – 290 с.
- Машинное обучение (курс лекций, К.В.Воронцов)  
[http://www.machinelearning.ru/wiki/index.php?title=Машинное обучение](http://www.machinelearning.ru/wiki/index.php?title=Машинное_обучение) 5.  
<https://www.machinelearningmastery.ru/>