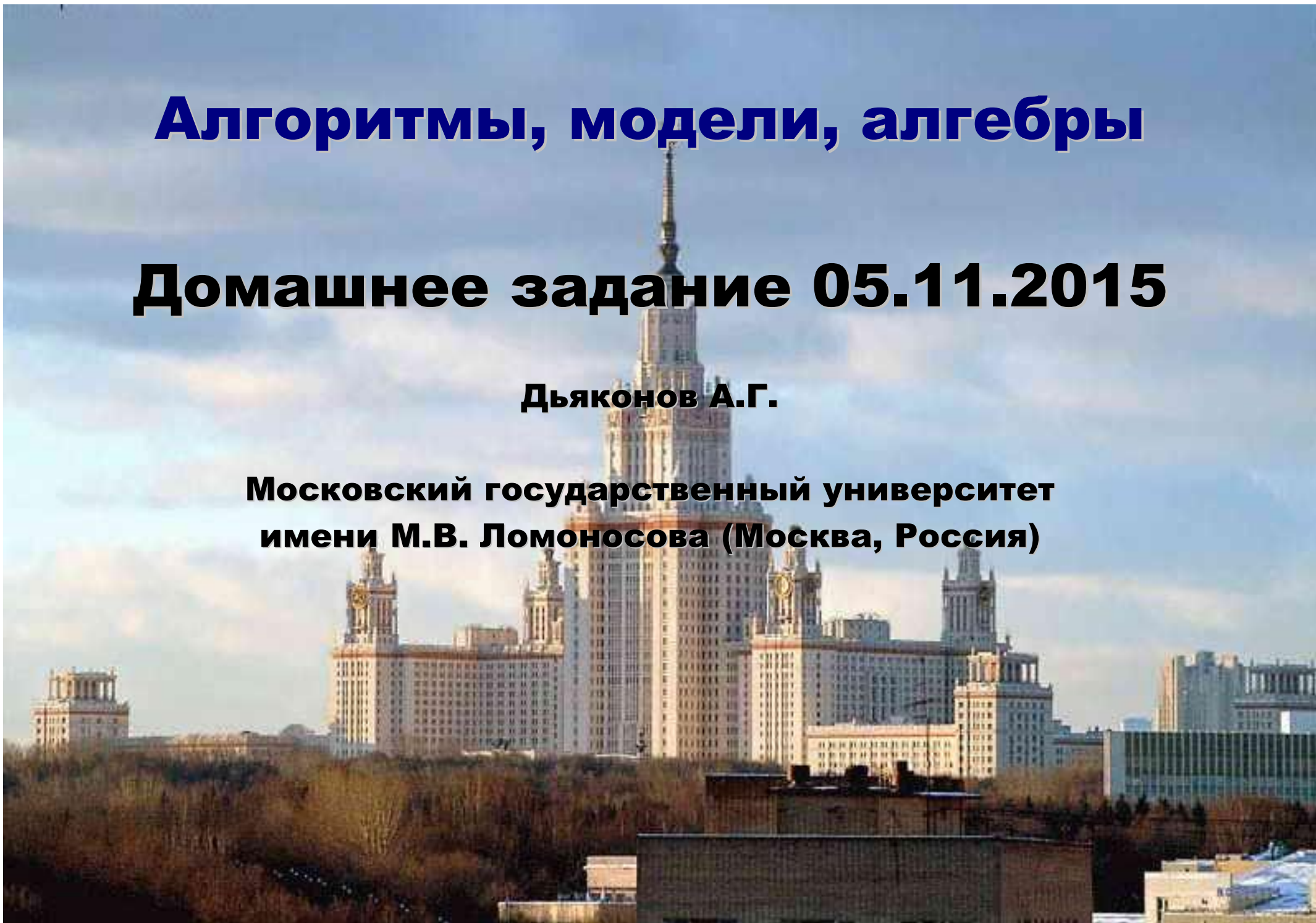


Алгоритмы, модели, алгебры

Домашнее задание 05.11.2015

Дьяконов А.Г.

**Московский государственный университет
имени М.В. Ломоносова (Москва, Россия)**



Домашнее задание

Выдано: 05.11.2015

Срок: до 18.11.2015, 23:59

В ветку форума

<https://inclass.kaggle.com/c/msuvisits/forums/t/17301/scikit-learn>

необходимо записать отчёт в виде pdf-презентации код.

Суть задания

- 1. Изучить алгоритмы библиотеки**
- 2. Провести с ними эксперименты:
научиться настраивать параметры**

Распределение по вариантам

1 Апишев Мурат Азаматович	Ridge Regression / Kernel ridge regression
2 Гой Антон Сергеевич	AdaBoost
3 Готман Мария Леонидовна	Lasso / LARS Lasso
4 Гурьянов Алексей Константинович	Linear and quadratic discriminant analysis
5 Жосан Юлия Сергеевна	Bayesian Ridge Regression
6 Кибитова Валерия Николаевна	Logistic regression
7 Козлов Владимир Дмитриевич	Stochastic Gradient Descent
8 Кузенко Татьяна Вячеславовна	Orthogonal Matching Pursuit (OMP)
9 Лукашкина Юлия Николаевна	Gradient Tree Boosting
10 Ожерельев Илья Сергеевич	Support Vector Machines
11 Сендерович Никита Леонидович	Semi-Supervised
12 Скробот Дмитрий Владиславович	Gaussian Processes
13 Темирчев Павел Георгиевич	Naive Bayes
14 Федосов Виктор Николаевич	Decision Trees
15 Хомутов Никита Юрьевич	Forests of randomized trees
16 Шапулин Андрей Валентинович	Probability calibration

Пояснение к заданию

- 1. Для каждой модели алгоритмов придумать несколько модельных задач**
 - на которых модель, видимо, работает хорошо
 - на которых она должна плохо работать

Задачи должны быть разной размерности (или параметризованы по размерности).

- 2. Если модель допускает, то надо провести исследования как в задаче регрессии, так и в задаче классификации (пояснить в чём различия с точки зрения настройки алгоритмов)**

- 3. Написать модуль настройки весов модели (а также проведения экспериментов)**

Провести эксперименты

**1. Как качество/скорость работы зависят от размеров задачи
(числа объектов/признаков)**

2. Как качество зависит от баланса классов/шумов/выбросов

**3. Формализовать понятие "качество"
(использовать различные функционалы)**

Требования к отчётам

**1. Описание модели
(чётко и грамотно объяснить)**

**2. Описание задач и экспериментов
(с обоснованием, почему выбрана такая схема экспериментов)**

3. Описание как пользоваться алгоритмом и как его настраивать

4. Выводы по результатам исследований

5. Самодостаточные рисунки и фрагменты кода