

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. ЛОМОНОСОВА

Факультет вычислительной математики и кибернетики

В. Ю. Королев, И. Г. Шевцова

Байесовские методы анализа больших данных

Учебное пособие



Москва – 2020

УДК 519.2(075.8)
ББК 22.17я73
К68

Рецензенты:

Ю. С. Хохлов – д-р физ.-мат. наук, проф. ВМК МГУ;
Б. Ф. Безродный – д-р техн. наук, зав. кафедрой прикладной математики МАДИ

Королев В. Ю., Шевцова И. Г.

К68 Байесовские методы анализа больших данных : учебное пособие / В. Ю. Королёв, И. Г. Шевцова. – Москва : издательский отдел ВМК МГУ имени М. В. Ломоносова; МАКС Пресс, 2020. – 140 с.

ISBN 978-5-89407-610-2 (ВМК МГУ имени М. В. Ломоносова)

ISBN 978-5-317-06504-1 (МАКС Пресс)

Учебное пособие содержит материал курсов лекций по байесовским методам анализа больших данных, которые читаются авторами на факультете ВМК и экономическом факультете МГУ. Пособие знакомит студентов с основными принципами выбора моделей, со смешанными моделями распределений, лежащими в основе байесовского подхода к анализу данных, и некоторыми алгоритмами анализа данных. В качестве примера применения байесовского подхода рассматривается математическая модель, лежащая в основе биологического паспорта спортсмена.

Пособие предназначено для студентов, обучающихся по магистерским программам кафедры математической статистики «Статистический анализ и прогнозирование рисков», «Интеллектуальный анализ больших данных» (совместно с кафедрой интеллектуальных информационных технологий) и «Анализ данных в экономике» (совместно с экономическим факультетом МГУ).

Ключевые слова: байесовские методы, интеллектуальный анализ данных, смешанное распределение, вероятностная модель, ЕМ-алгоритм.

УДК 519.2(075.8)
ББК 22.17я73

ISBN 978-5-89407-610-2
ISBN 978-5-317-06504-1

© Королев Ю. В., Шевцова И. Г., 2020
© Издательский отдел ВМК МГУ, 2020
© ООО «МАКС Пресс», 2020

Korolev Victor, Shevtsova Irina.

Bayesian Methods for Big Data Mining : tutorial / Irina Shevtsova, Victor Korolev. – Moscow : CMC MSU; MAKS Press, 2020. – 140 p.

ISBN 978-5-89407-610-2 (CMC MSU)

ISBN 978-5-317-06504-1 (MAKS Press)

The book contains the material of lectures in Bayesian methods for big data mining read by the authors at the Faculty of Computational Mathematics and Cybernetics and the Faculty of Economics, Lomonosov Moscow State University. The book introduces the students into principles of formation of mathematical models, mixed probability models underlying the Bayesian approach to data mining as well as into some algorithms of data analysis. As an example, the basic mathematical model used for the construction of the biological passport of a sportsman is considered.

The book is oriented for the students of the master programs of the Department of Mathematical Statistics “Statistical Analysis and Risk Prediction”, “Big Data Mining” (in cooperation with the Department of Intellectual Information Technologies) and “Data Analysis in Economics” (in cooperation with the Faculty of Economics, Lomonosov Moscow State University).

Key words: Bayesian methods, data mining, mixed distribution, probability model, EM-algorithm.

Учебное издание

КОРОЛЕВ Виктор Юрьевич
ШЕВЦОВА Ирина Геннадьевна

**БАЙЕСОВСКИЕ МЕТОДЫ АНАЛИЗА
БОЛЬШИХ ДАННЫХ**

Учебное пособие

Издательский отдел
Факультета вычислительной математики и кибернетики
МГУ имени М.В. Ломоносова
Лицензия ИД N 05899 от 24.09.01 г.
119992, ГСП-2, Москва, Ленинские горы,
МГУ имени М.В. Ломоносова,
2-й учебный корпус

Издательство «МАКС Пресс»
Главный редактор: *Е. М. Бугачева*

Напечатано с готового оригинал-макета
Подписано в печать 27.11.2020 г.
Формат 70х100 1/16. Усл.печ.л. 11,4.
Тираж 100 экз. Заказ 181.

Издательство ООО «МАКС Пресс»
Лицензия ИД N 00510 от 01.12.99 г.
119992, ГСП-2, Москва, Ленинские горы,
МГУ им. М.В. Ломоносова,
2-й учебный корпус, 527 к.
Тел. 8(495)939-3890/91. Тел./Факс 8(495)939-3891.

Отпечатано в полном соответствии с качеством
предоставленных материалов в ООО «Фотоэксперт»
115201, г. Москва, ул. Котляковская, д.3, стр. 13.