

Савченко В.А.

Научный руководитель: к. т. н., доц. Рыжкова М.Н.

Муромский институт (филиал) федерального государственного образовательного учреждения высшего образования «Владимирский государственный университет

имени Александра Григорьевича и Николая Григорьевича Столетовых»

602264, г. Муром, Владимирская обл., ул. Орловская, 23

E-mail: vladislavasavcenko@gmail.com

Разработка функциональной модели системы построения аппроксимирующего полинома для обработки экспериментальных данных

Данные, получаемые в ходе выполнения какого-либо эксперимента, обычно являются дискретным представлением некой функциональной зависимости. Такой вид представления чаще всего затрудняет использование этих данных в математическом моделировании, поэтому почти каждое научное исследование использует теорию аппроксимации. Аналитическое представление экспериментальных данных необходимо и важно, особенно для обеспечения высокой точности проводимых исследований.

Аппроксимация данных является одной из задач регрессионного анализа, который стремительно развивается в машинном обучении. Преимущество данного подхода заключается в получении единственной модели, которая полностью опишет множественные взаимные связи между входными и выходными данными, даже в случае нелинейности связи между ними [1].

В ходе выполнения эксперимента может быть получено огромное количество объектов исследования, которые невозможно обработать вручную. Как правило, это большие таблицы, в которых есть несколько нецелевых признаков и один целевой. Данные, получаемые в ходе выполнения эксперимента, имеют разный масштаб, пропуски и аномалии, поэтому перед выполнением задачи аппроксимации необходимо выполнить предварительную обработку, чтобы структурировать их. Разработка функциональной модели позволит грамотно и последовательно описать работу выполнения алгоритмов системой [2].

Задача полиномиальной аппроксимации сводится к решению нормированной системы линейных уравнений. При больших степенях такая система является плохо обусловленной, хотя и имеет единственное решение, но чаще всего смысла оно не имеет. Поэтому стоит рассмотреть наивысшую степень полинома, при которой решение данной задачи будет единственно верным. Для полиномиальной аппроксимации оптимальной считается степень не выше пятой [3].

Целью данного исследования является разработка функциональной модели для обобщения работы системы построения аппроксимирующего полинома.

Для достижения поставленной цели требуется выполнить следующие задачи:

1. разработать модель «чёрный ящик» для определения вида входных и выходных параметров;
2. разработать модель состава для определения основных блоков обработки данных;
3. разработать структурную модель для определения перемещения данных в блоках;
4. разработать функциональную модель для определения работы отдельных блоков системы.

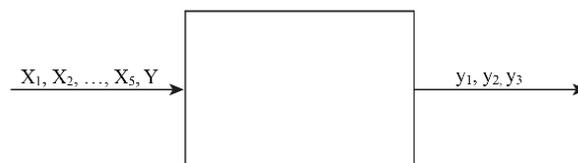


Рис. 1. Модель «чёрный ящик»

На вход системе подаётся таблица экспериментальных данных, которая может содержать 1-5 нецелевых факторов X и одно значение отклика Y . Выходными данными системы являются: y_1

– аппроксимирующее уравнение в виде полинома 5-ой степени; y_2 – значение точности работы системы в процентах, y_3 – рекомендация для пользователя.

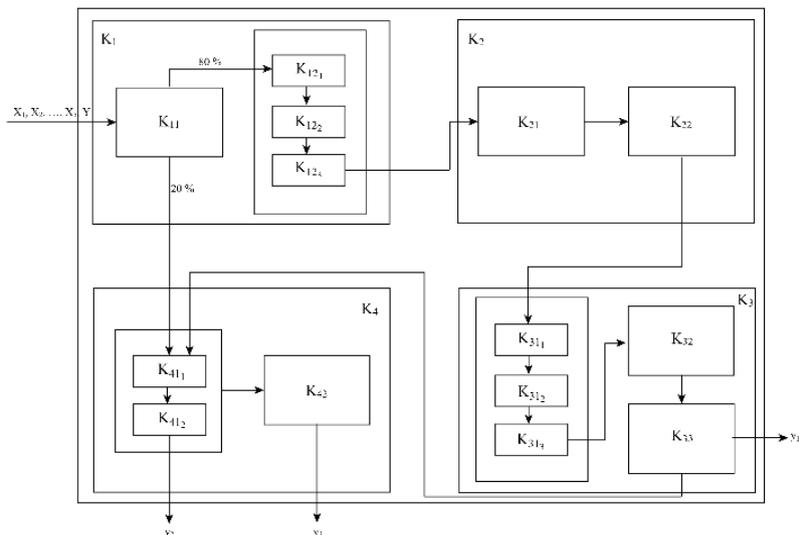


Рис. 2. Функциональная модель

- K_1 – блок предварительной обработки данных, который содержит в себе:
- K_{11} – разбиение входного набора на обучающую (80 %) и тестовую (20 %) выборки;
 - K_{12} – предварительная обработка данных, а именно нормирование, заполнение пропусков и поиск выбросов в обучающей выборке.
- K_2 – блок отбора признаков:
- K_{21} – вычисление матрицы корреляции;
 - K_{22} – поиск линейной зависимости между нецелевыми факторами и откликом.
- K_3 – блок вычисления коэффициентов аппроксимирующего уравнения:
- K_{31} – применение метода наименьших квадратов для вычисления коэффициентов аппроксимирующего уравнения, который заключается в составлении системы уравнений, её нормировании и решении методом Гаусса;
 - K_{32} – переход от нормированных коэффициентов и вычисление свободного члена;
 - K_{33} – составление аппроксимирующего уравнения.
- K_4 – блок тестирования, полученной математической модели:
- K_{41} – вычисление точности, которое заключается в подстановке нецелевых факторов из тестовой выборки в полученную математическую модель для вычисления теоретического значения отклика, которое после сравнивается с экспериментальным значением отклика;
 - K_{42} – оценка математической модели, а именно вывод рекомендации для пользователя если значение точности менее 90 %.

В ходе исследования был выполнен ряд поставленных задач, которые обеспечили разработку функциональной модели, описывающей последовательность работы системы и перемещение данных в ней. Перспективой дальнейшего исследования является реализация системы построения аппроксимирующего полинома для обработки экспериментальных данных.

Литература

1. Шашков В.Б. Прикладной регрессионный анализ. Многофакторная регрессия: Учебное пособие.-Оренбург: ГОУ ВПО ОГУ, 2003. – 363с.
2. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных / пер. с англ. А. А. Слинкина. – М.: ДМК Пресс, 2015. – 400 с.: ил.
3. Амосов А.А., Дубинский Ю.А., Копченова Н.В. Вычислительные методы для инженеров: Учеб. пособие. — М.:Выш. шк., 1994. — 544 с.