



РЕГРЕССИОННЫЙ МЕТОД В ИНФОКОММУНИКАЦИОННОЙ ТЕХНОЛОГИИ ПРИ ИССЛЕДОВАНИИ ЭФФЕКТИВНОСТИ ПРИМЕНЕНИЯ СПУТНИКОВЫХ ДАННЫХ

Рокина Ирина Константиновна

магистрант, Поволжский Государственный технологически университет,
РФ, г. Йошкар-Ола

Из года в год на всей территории земного шара преобладает проблема лесных пожаров, неконтролируемого стихийного бедствия, пагубно отражающегося на экономике республики в целом и лесном комплексе в частности. Для решения такой масштабной задачи необходимо использовать различные средства и методы системы раннего обнаружения лесных пожаров.

Одним из методов анализа данных является регрессионный анализ, который позволяет определить связь между зависимой переменной и набором независимых, и использовать эту связь для прогнозирования значений.

Наиболее распространённый вид регрессионного анализа — линейная регрессия, когда находят линейную функцию, которая, согласно определённым математическим критериям, наиболее соответствует данным. Например, в методе наименьших квадратов вычисляется прямая (или гиперплоскость), сумма квадратов между которой и данными минимальна.

Математическое определение регрессии можно определить следующим образом. Пусть Y - зависимая переменная, а $x_1, x_2, \dots, x_n$ независимые переменные (факторы), определяющие поведение зависимой переменной. При построении модели, описывающей зависимость Y от $x_1, x_2, \dots, x_n$, предполагается, во-первых, что у исследователя имеются результаты совокупных наблюдений зависимой переменной Y и независимых переменных $x_1, x_2, \dots, x_n$, во-вторых, что значения независимых переменных определяются точно (без ошибок), а значение зависимой переменной Y определяется с ошибками, имеющими случайный характер. Математическая модель, описывающая данные такого вида, выглядит следующим образом:

$$Y = f(x_1, x_2, \dots, x_n) + \varepsilon$$

где: ε - случайная ошибка наблюдений зависимой переменной.

Таким образом, регрессия описывает поведение наблюдаемой зависимой переменной в среднем, представляя ее главную тенденцию. В связи с этим нахождение регрессии по результатам наблюдений называют сглаживанием данных.

Исходя из географических особенностей исследуемой территории общая площадь лесного фонда составляет 90% от всей площади территории. Из них сосна занимает большую часть – 41 %.

По проведенному анализу статистических данных, взятых из ГИС было выявлено, что на территории одного муниципального образования за последние 5 лет произошло наибольшее количество возгораний. Анализ горимости лесов показал, что одной из основных причин возникновения пожаров, являлось неосторожное обращение с огнем местного населения. Возникновению пожаров способствует не только человеческий фактор, но и такие как климат, погода, влияющая на созревание лесных горючих материалов (ЛГМ), пирологические характеристики растительности, антропогенное воздействие [1, 2, 3].

В случае системы раннего обнаружения лесных пожаров, зависимой переменной может быть вероятность возникновения пожара, а независимыми переменными – метеорологические показатели. Для построения регрессионной модели использовались статистические данные о значениях этих переменных в периоды, когда происходили пожары. Такая модель будет полезна для предсказания вероятности возникновения пожара на основе текущих значений переменных, а также для своевременного (оперативного) принятия мер по борьбе с возгораниями.

Так как в модели используется несколько факторов влияющих на количество возгораний, то будем применять множественную линейную регрессию.

Формула множественной линейной регрессии имеет вид:

$$Y = R_0 + R_1 * X_1 + R_2 * X_2 + \dots + R_n * X_n$$

где: Y – количество возгораний (термические точки);

X_1 - температура воздуха (градусы Цельсия) на высоте 2 метра над поверхностью земли;

X_2 - скорость ветра на высоте 10-12 метров над земной поверхностью (метры в секунду);

X_3 - относительная влажность (%) на высоте 2 метра над поверхностью земли;

R_0 - свободный член;

R_1, R_2, R_3 - коэффициенты регрессии.