Автор:

Дмитриев Егор Андреевич

студент

ФГАОУ ВО «Самарский национальный исследовательский университет им. академика С.П. Королева»

г. Самара, Самарская область

МЕТОД ОПОРНЫХ ВЕКТОРОВ

Аннотация: *в статье рассматривается один из методов машинного обучения – использование метода опорных векторов.*

Ключевые слова: *Линейный классификатор, алгоритм классификации, функционал.*

Основные определения

Определение 1. Линейный классификатор – алгоритм классификации, который основывается на применении разделяющей поверхности. Разделяющая поверхность называется гиперплоскостью, подпространство размерности n-1 пространства признаков.

Определение 2. Алгоритм классификации – отображение $a : X \to Y$, где X – метрическое пространство признаков классифицируемого объекта, а Y – множество классов. Классифицировать объект – однозначно определить, к какому классу относится объект.

Определение 3. Функционал – отображение $a : X \to R$, где X – метрическое пространство признаков классифицируемого объекта, а R – множество действительных чисел.

Введение

Машинная обучение – это сфера, которая является смежной с методами оптимизации и статистикой. Как и в методах оптимизации в машинном обучении нужно найти оптимальные параметры, для минимизации какого-то функционала. Метод опорных векторов является широко распространенным методом, который используется в методах оптимизации, а также и в машинном обучении.

Постановка проблемы

Необходимо ознакомиться с принципами работы метода опорных векторов.

Описание метода опорных векторов

Суть задачи классификации заключается в построении алгоритма классификации по обучаемой выборке. Идея линейных классификаторов заключается в построении линейных функций, являющимися разделяющими поверхности, по обе стороны которой лежат объекты разных классов. При использовании метода опорных векторов используется функционал особого вида:

$$\sum_{i=1}^{l}(1-M_i(w,w_0))_+ + \frac{1}{2C}||w||^2 \to min_{w,w_0},$$

где $M_i(w,w_0) = y_i((w,x_i) - w_0)$ – величина отступа.

Пусть имеем объекты, принадлежащие двум классам. Если выборка является линейно разделимой, то оптимальную разделяющую поверхность можно найти из решения системы:

$$\begin{cases} ||w||^2 \to min \\ M_i(w,w_0) \geq 1 \end{cases} i = 1, \dots, l$$

С учетом того, что выборка на практике чаще всего не является линейно разделимой, ограничение в системе становится менее жестким:

$$\begin{cases} \frac{1}{2}||w||^2 + C\sum_{i=1}^{l}\xi_i \to min_{w,w_0,\xi} \\ M_i(w,w_0) \geq 1 - \xi_i, i = 1, \dots, l \\ \xi_i \geq 0, i = 1, \dots, l \end{cases}$$

Имеем задачу математического программирования. Целевая функция является непрерывной и определена на компактном множестве, следовательно, по теореме Вейерштрасса целевая функция достигает минимума. Изменяя ограничения, получаем задачу, к которой применяем условия Каруша – Куна – Таккера для нахождения минимума функционала:

$$\varphi(w,w_0,\xi,\lambda,\eta) = \frac{1}{2}||w||^2 - \sum_{i=1}^{l}\lambda_i(M_i(w,w_0) - 1) - \sum_{i=1}^{l}\xi_i(\lambda_i + \eta_i - C)$$

$$\begin{cases} \dfrac{\partial \varphi}{\partial w} = 0, \dfrac{\partial \varphi}{\partial w_0} = 0, \dfrac{\partial \varphi}{\partial \lambda} \\ \xi_i \geq 0, \lambda_i \geq 0, \eta_i \geq 0 \\ \lambda_i = 0 \text{ либо } M_i(w, w_0) = 1 - \xi_i \end{cases}$$

При решении системы в зависимости от параметров получаются 3 типа объектов:

1) периферийные, неинформативные объекты;

2) опорные;

3) объекты, которые оказались по другую сторону от разделяющей поверхности.

Заключение

В данной работе был рассмотрен принципы работы метода опорных векторов.

Список литературы

1. Айвазян С.А. Прикладная статистика: классификация и снижение размерности / С.А. Айвазян, В.М. Бухштабер, И.С. Енюков, Л.Д. Мешалкин. – М.: Финансы и статистика, 1989.

2. Загоруйко Н.Г. Прикладные методы анализа данных и знаний. – Новосибирск: ИМ СО РАН, 1999.

3. Вапник В.Н. Восстановление зависимостей по эмпирическим данным. – М.: Наука, 1979.