Data mining - zadanie 24

Gabriel Budziński

November 13, 2023

1 Treść

Explain how we may get the multinomial (multi-class) logistic model for K classes by running $K-1$ independent binary logistic regression models.

2 Przypomnienie

W regresji logistycznej mamy do czynienia z dwiema kategoriami (0 i 1) i mamy dwa prawdopodobieństwa: wynik x jest w kategorii 1 z prawdopodobieństwem $p(x)$ oraz w kategorii 0 z prawdopodobieństwem $1-p(x)$. Znamy też funkcję $\text{logit} p(x) = ln\left(\frac{p(x)}{1-p(x)}\right) = \beta_0 + \beta_1 x$.

3 Rozwiązanie

Aby otrzymać model multiklasowy (więcej niż dwa możliwe rezultaty) możemy $K-1$ razy niezależnie przeprowadzić binarną regresję logistyczną, gdzie jeden z wyników (kategorii) będzie *pivotem*, a pozostałe $K-1$ wyników są osobno regresowane w oparciu o *pivot*. Jeśli jako *pivot* wybraliśmy ostatni, K-ty rezultat, to równania dla pozostałych mają postać (*ten wzór jest znany też jako Additive Log Ratio*).

$$\ln \frac{\Pr(Y_i = k)}{\Pr(Y_i = K)} = \beta_k \cdot \mathbf{X}_i, \quad k < K$$

gdzie $\Pr(Y_i = k)$ - prawdopodobieństwo, że i-ta obserwacja jest w klasie k, β_k - wektor współczynników regresji dla klasy k, $\mathbf{X}_i$ - wektor *explanatory variables* dla i-tej obserwacji

Po nałożeniu funkcji exp na obie strony otrzymujemy

$$\Pr(Y_i = k) = \Pr(Y_i = K) e^{\beta_k \cdot \mathbf{X}_i}$$

Jako że suma prawdopodobieństw po wszystkich K kategoriach to 1 mamy

$$\Pr(Y_i = K) = 1 - \sum_{j=1}^{K-1} \Pr(Y_i = j) = 1 - \sum_{j=1}^{K-1} \Pr(Y_i = K) e^{\beta_j \cdot \mathbf{X}_i} \implies$$

$$\Pr(Y_i = K) = \frac{1}{1 + \sum_{j=1}^{K-1} e^{\beta_j \cdot \mathbf{X}_i}}$$

Z uprzedniej równości

$$\Pr(Y_i = k) = \frac{e^{\beta_k \cdot \mathbf{X}_i}}{1 + \sum_{j=1}^{K-1} e^{\beta_j \cdot \mathbf{X}_i}}$$

W ten sposób otrzymaliśmy prawdopodobieństwa bycia w danej kategorii dla zmiennej $\mathbf{X}_i$.

Uwaga: model opiera się na założeniu o niezależności nieistotnych alternatyw, bo przeprowadzamy wielokrotne regresje.