Interpretable Machine Learning

Generalized Linear Models



 $\pi(\mathbf{x})$

Learning goals

- Definition of GLMs
- Logistic regression as example
- Interpretation in logistic regression



GENERALIZED LINEAR MODEL (GLM) • Nelder and Wedderburn 1972

Problem: Target variable given feat. not always normally dist. ~> LM not suitable

- Target is binary (e.g., disease classification)
 → Bernoulli / Binomial distribution
- Target is count variable (e.g., number of sold products)
 Poisson distribution
- Time until an event occurs (e.g., time until death)
 → Gamma distribution





GENERALIZED LINEAR MODEL (GLM) • Nelder and Wedderburn 1972

Problem: Target variable given feat. not always normally dist. \rightsquigarrow LM not suitable

- Target is binary (e.g., disease classification) → Bernoulli / Binomial distribution
- Target is count variable (e.g., number of sold products)
 Poisson distribution
- Time until an event occurs (e.g., time until death)
 → Gamma distribution





Solution: GLMs - extend LMs by allowing other distributions from exponential family

$$g(\mathbb{E}(y \mid \mathbf{x})) = \mathbf{x}^{\top} \theta \iff \mathbb{E}(y \mid \mathbf{x}) = g^{-1}(\mathbf{x}^{\top} \theta)$$

- Link function *g* links linear predictor **x**^Tθ to expectation of distribution of *y* | **x** → LM is special case: Gaussian distribution for *y* | **x** with *g* as identity function
- Link function g and distribution need to be specified
- High-order and interaction effects can be manually added as in LMs
- Note: Interpretation of weights depend on link function and distribution

GLM - LOGISTIC REGRESSION

• Logistic regression $\hat{=}$ GLM with Bernoulli distribution and logit link function:

$$g(x) = \log\left(\frac{x}{1-x}\right) \Rightarrow g^{-1}(x) \qquad \qquad = \frac{1}{1+\exp(-x)}$$

• Models probabilities for binary classification by

$$\pi(\mathbf{x}) = \mathbb{E}(y \mid \mathbf{x}) = P(y = 1) = g^{-1}(\mathbf{x}^{\top}\theta) = \frac{1}{1 + \exp(-\mathbf{x}^{\top}\theta)}$$





GLM - LOGISTIC REGRESSION

- Typically, we set the threshold to 0.5 to predict classes, e.g.,
 - Class 1 if π(x) > 0.5
 - Class 0 if $\pi(\mathbf{x}) \leq 0.5$





GLM - LOGISTIC REGRESSION - INTERPRETATION

- Recall: Odds is ratio of two probabilities, odds ratio compares ratio of two odds
- Weights θ_j are interpreted linear as in LM (but w.r.t. log-odds)
 → difficult to comprehend

$$log-odds = \log\left(\frac{\pi(\mathbf{x})}{1-\pi(\mathbf{x})}\right) = \log\left(\frac{P(y=1)}{P(y=0)}\right) = \theta_0 + \theta_1 x_1 + \ldots + \theta_p x_p$$

Interpretation:

Changing x_i by one unit, changes log-odds of class 1 compared to class 0 by θ_i



GLM - LOGISTIC REGRESSION - INTERPRETATION

- Recall: Odds is ratio of two probabilities, odds ratio compares ratio of two odds
- Weights θ_j are interpreted linear as in LM (but w.r.t. log-odds)
 → difficult to comprehend

$$log-odds = \log\left(\frac{\pi(\mathbf{x})}{1-\pi(\mathbf{x})}\right) = \log\left(\frac{P(y=1)}{P(y=0)}\right) = \theta_0 + \theta_1 x_1 + \ldots + \theta_p x_p$$

Interpretation:

Changing x_j by one unit, changes log-odds of class 1 compared to class 0 by θ_j

- Odds for class 1 vs. class 0: $odds = \frac{\pi(\mathbf{x})}{1 \pi(\mathbf{x})} = \exp(\theta_0 + \theta_1 x_1 + \ldots + \theta_p x_p)$
- Instead of interpreting changes w.r.t. log-odds, odds ratio is more common

$$=\frac{odds_{x_{j+1}}}{odds}=\frac{\exp(\theta_0+\theta_1x_1+\ldots+\theta_j(x_j+1)+\ldots+\theta_px_p)}{\exp(\theta_0+\theta_1x_1+\ldots+\theta_jx_j+\ldots+\theta_px_p)}=\exp(\theta_j)$$

Interpretation: Changing x_j by one unit, changes the **odds ratio** for class 1 (compared to class 0) by the **factor** $\exp(\theta_j)$



GLM - LOGISTIC REGRESSION - EXAMPLE

- Create a binary target variable for bike rental data:
 - $\bullet\,$ Class 1: "high number of bike rentals" >70% quantile (i.e., ${\rm cnt}>5531)$
 - $\bullet\,$ Class 0: "low to medium number of bike rentals" (i.e., ${\rm cnt} \leq 5531)$
- Fit a logistic regression model (GLM with Bernoulli distribution and logit link)

	Weights	SE	p-value
(Intercept)	-8.52	1.21	0.00
seasonSPRING	1.74	0.60	0.00
seasonSUMMER	-0.86	0.77	0.26
seasonFALL	-0.64	0.55	0.25
temp	0.29	0.04	0.00
hum	-0.06	0.01	0.00
windspeed	-0.09	0.03	0.00
days_since_2011	0.02	0.00	0.00



GLM - LOGISTIC REGRESSION - EXAMPLE

- Create a binary target variable for bike rental data:
 - Class 1: "high number of bike rentals" > 70% quantile (i.e., cnt > 5531)
 - Class 0: "low to medium number of bike rentals" (i.e., $\mathrm{cnt} \leq 5531$)
- Fit a logistic regression model (GLM with Bernoulli distribution and logit link)

	Weights	SE	p-value	1.00	
(Intercept)	-8.52	1.21	0.00	e re	
seasonSPRING	1.74	0.60	0.00	of big	
seasonSUMMER	-0.86	0.77	0.26	Ser effect	
seasonFALL	-0.64	0.55	0.25	o nun	
temp	0.29	0.04	0.00	. high	
hum	-0.06	0.01	0.00	t ss	
windspeed	-0.09	0.03	0.00	5 %	
days_since_2011	0.02	0.00	0.00	-10	0 10 20 30 Temperature in °C

Interpretation

 If temp increases by 1°C, odds ratio for class 1 increases by factor exp(0.29) = 1.34 compared to class 0, c.p. (≙ "high number of bike rentals" now 1.34 times more likely)