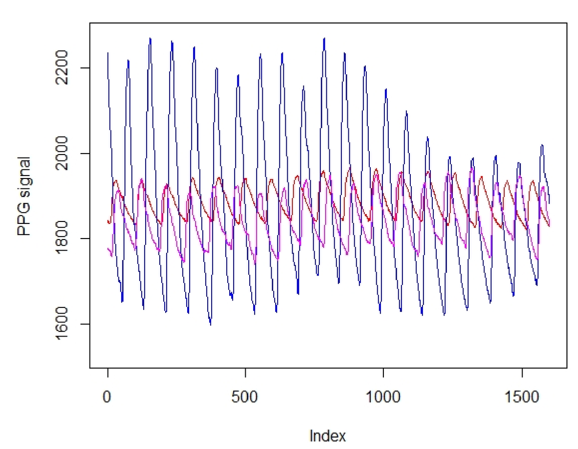
CNN COMBINED WITH MIXED LOGISTIC REGRESSION TO DEAL WITH SIGNALS FOR CLASSIFICATION

Chen Xiao

†Department of Mathematics, Hong Kong Baptist University

Introduction

- DATA:
  1.N participants, each have several measurements at a different time to get a 2-class label.
  2. Signal: sequence length is s, d channels.
  3.K features such as age, gender, weight ,etc.
  TARGET:
  Given signals, we want to know which of the two classes the participants belongs to at a given moment.
- Sample case:
  1.PPG signal: related to blood pressure, can be continuously monitored, one channel, 50Hz 32s
  2.576 participants,6022 records
  3.Class label: high blood pressure, normal(Unbalanced)
  4.Features such as age, gender, weight, etc.
  5.Task:Detection of high or normal blood pressure based on PPG signal



Model

- S_{ij} : Signals:d channels
- x_{ij}:K Features such as age, gender, weight ,etc.
- $$\eta_{ij} = \log\frac{\pi_{ij}}{1-\pi_{ij}} = x_{ij}^T\beta + \gamma f_\theta(S_{ij}) + u_i,$$
  $$Y_{ij} \sim Bernoulli(\pi_{ij})$$
  where random effects $u_i \sim \mathcal{N}(0,\sigma^2)$.
  N individuals
  Each individuals have J_i records.
- $f_\theta(S_{ij})$: CNN
  Inputs: signals N, sequence length , channels
  Output: One dimension $f_\theta(S_{ij})$
- $b = (\beta^T, \gamma)^T$

$$\begin{aligned}&P(Y_{ij}|x_{ij}, S_{ij}; b, \theta, u_i)\\ &= \exp(Y_{ij}\eta_{ij})\sigma(-\eta_{ij})\\ &\geq \exp(Y_{ij}\eta_{ij})\sigma(\epsilon_{ij})\exp[-\lambda(\epsilon_{ij})(\eta_{ij}^2 - \epsilon_{ij}^2) - \frac{\eta_{ij}+\epsilon_{ij}}{2}]\\ &= h(Y_{ij}|x_{ij}, S_{ij}; b, \theta, u_i)\end{aligned}$$

According to the form of $h(Y_{ij}|x_{ij}, S_{ij}; b, \theta, u_i)$ We conclude that the lower bound of the posterior of u_i is proportional to a normal distribution.

Model

$$\prod_{j=1}^{J_i} h(Y_{ij}|x_{ij}, S_{ij}; b, \theta, u_i)p(u_i|\sigma^2) \propto N(u_i; \mu_i, \sigma_i^2),$$

where

$$\mu_i = \sigma_i^2 \sum_{j=1}^{J_i}(Y_{ij} - 2\lambda(\epsilon_{ij})(x_{ij}^T\beta + \gamma f_\theta(S_{ij})))$$

$$\sigma_i^2 = (\frac{1}{\sigma^2} + 2\sum_{j=1}^{J_i}\lambda(\epsilon_{ij}))^{-1}$$

The lower bound of the likelihood is

$$\begin{aligned}&\sum_{i=1}^{N}\log(Y_i|b,\sigma^2)\\ &\geq \sum_{i=1}^{N}\log\int f(Y_i, u_i|b,\theta,\sigma^2)du_i\\ &\geq \sum_{i=1}^{N}\int q(u_i)\log\frac{f(Y_i,u_i|b,\theta,\sigma^2)}{q(u_i)}du_i\\ &= \mathbb{E}_q\log f(Y_i,u_i|b,\theta,\sigma^2) - \mathbb{E}_q\log q(u_i)\end{aligned}$$

where $f(Y_i, u_i|b, \theta, \sigma^2) = \prod_{j=1}^{J_i} h(Y_{ij}|x_{ij}, S_{ij}; b, \theta, u_i)p(u_i|\sigma^2)$

Algorithm

1.In the E step, we get the posterior of u_i,

$$q(u_i|b^{old}, \epsilon_{ij}^{old}, \sigma^{2old}) = N(u_i; \mu_i, \sigma_i^2)$$

2.In the M step, we need to maximize $\mathbb{E}_q\log f(Y_i, u_i|b, \theta, \sigma^2) - -\mathbb{E}_q\log q(u_i)$
2.1 The updating equation of ϵ_{ij}^2 is

$$\epsilon_{ij}^2 = (x_{ij}^T\beta^{old})^2 + 2\gamma f_\theta(S_{ij}) + 2x_{ij}^T\beta^{old}u_i + (\mu_i^2 + \sigma_i^2).$$

2.2 Then we update the parameter θ in the CNN structure.
2.3The updating equation of β is

$$\beta = (\sum_{i=1}^{N}\sum_{j=1}^{J_i}\lambda(\epsilon_{ij})x_{ij}x_{ij}^T)^{-1}(\sum_{i=1}^{N}\sum_{j=1}^{J_i}(Y_{ij} - 0.5 - 2\lambda(\epsilon_{ij})\mu_i)x_{ij})$$

2.4 The updating equation of σ^2 is

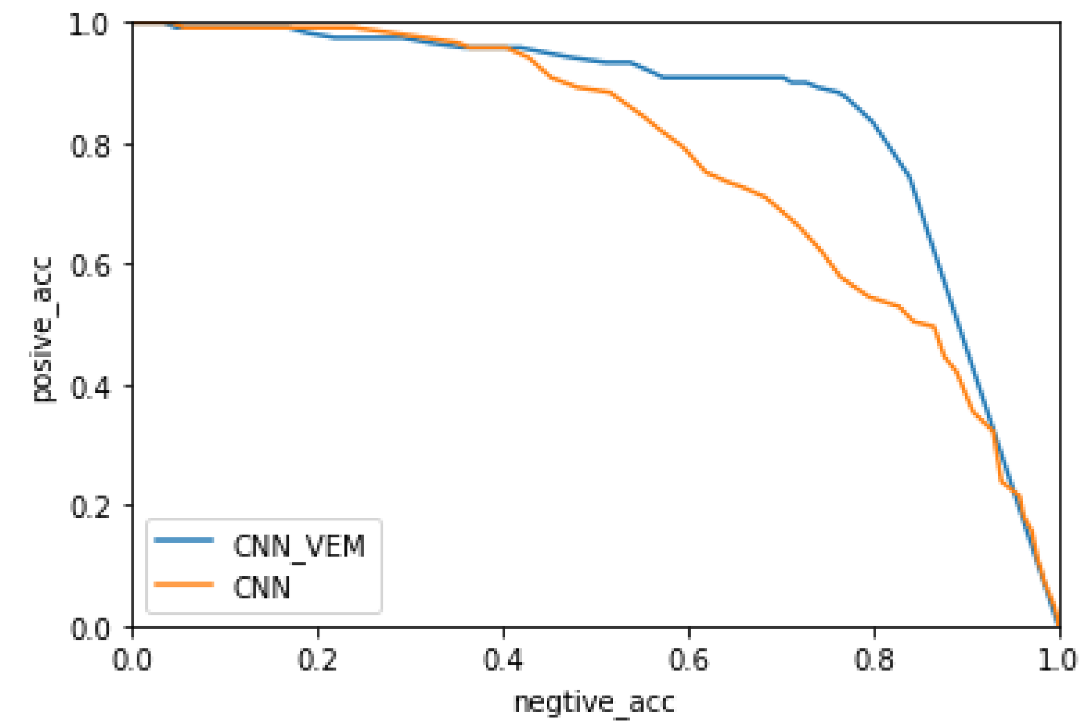
$$\sigma^2 = \frac{1}{N}(\sum_{i=1}^{N}(\mu_i^2 + \sigma_i^2))$$

References

[1] Rijmen, Frank, Vomlel, Jiri. (2008). *Journal of Statistical Computation and Simulation*, **78**(8), 765-779.

PPG and Blood pressure

The ROC curve



Another example:Human Activity Recognition

- 2 classes of activities: SITTING, STANDING
- The length of the sequence in time-series :128
- Number of channels where measurements are made : 9.
- 21 participants
- There are 9 channels in this case, which include 3 different accele measurements for each 3 coordinate axes.
- Training data size:2128,Testing data size: 532