



Equipe 22, UMRS 1138 : Sciences de l'information au service de la médecine personnalisée

Apprentissage supervisé pour la médecine :

MODÈLE PÉNALISÉ PRONOSTIQUE À VARIABLE LATENTE POUR DES DONNÉES CENSURÉES

application : prédiction du risque de réhospitalisation

par

Anne-Sophie Jannot et Simon Bussy

Présentation MAP5

13 mai 2016

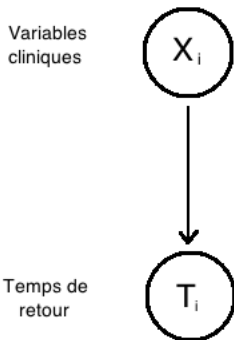
Objectif : Identifier les caractéristiques permettant de prédire le pronostic des patients

- Exemple 1 : quelles sont les caractéristiques qui font qu'un patient va répondre ou non à un traitement ?
Objectif : Proposer un traitement personnalisé

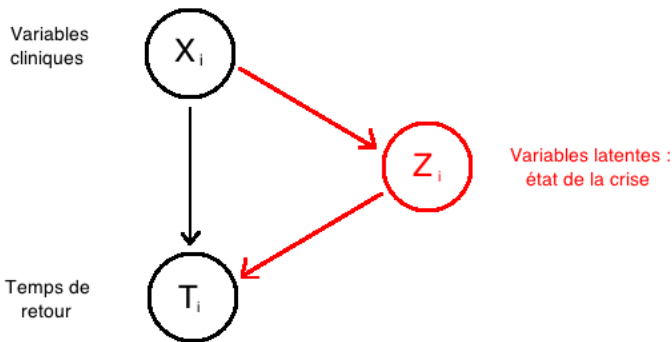
Objectif : Identifier les caractéristiques permettant de prédire le pronostic des patients

- Exemple 1 : quelles sont les caractéristiques qui font qu'un patient va répondre ou non à un traitement ?
Objectif : Proposer un traitement personnalisé
- Exemple 2 : quelles sont les caractéristiques liées au soin qui vont conduire un patient à être réhospitalisé précocément ? Objectif : éviter des réhospitalisations

Objectif : Identifier les caractéristiques permettant de prédire le pronostic des patients



Objectif : Identifier les caractéristiques permettant de prédire le pronostic des patients



Pourquoi un modèle pénalisé ?

- Réutilisation de données produites dans le cadre du soin. Données "en vie réelle". Nombreuses variables. Pertinence des données recueillies pour prédire le pronostic ?

Pourquoi un modèle pénalisé ?

- Réutilisation de données produites dans le cadre du soin. Données "en vie réelle". Nombreuses variables. Pertinence des données recueillies pour prédire le pronostic ?
- Variables longitudinales (médicaments, valeurs de laboratoires, monitoring...)

Pourquoi un modèle à variable latente ?

- La réponse au traitement ou aux soins est mesurée indirectement via le temps de réhospitalisation, récurrence ou décès

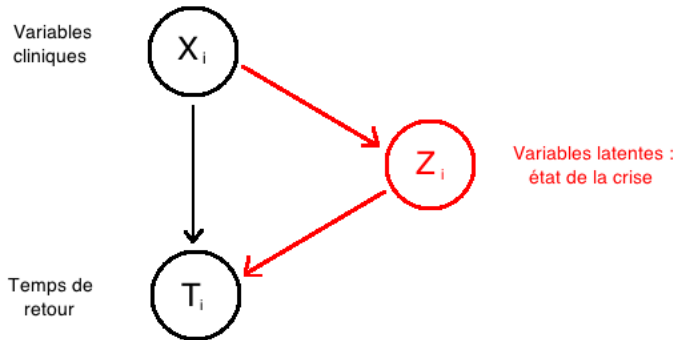
Pourquoi un modèle à variable latente ?

- La réponse au traitement ou aux soins est mesurée indirectement via le temps de réhospitalisation, récurrence ou décès
- Exemple 1 : réponse à une chimiothérapie = temps nécessaire pour réduire la taille de la tumeur de XXX

Pourquoi un modèle à variable latente ?

- La réponse au traitement ou aux soins est mesurée indirectement via le temps de réhospitalisation, récurrence ou décès
- Exemple 1 : réponse à une chimiothérapie = temps nécessaire pour réduire la taille de la tumeur de XXX
- Exemple 2 : prise en charge hospitalière adéquate = absence de réhospitalisation précoce = temps jusqu'à la prochaine hospitalisation

Pourquoi un modèle à variable latente ?



Pourquoi un modèle pour des données censurées ?

- Suivi limité dans le temps

Pourquoi un modèle pour des données censurées ?

- Suivi limité dans le temps
- Censure : date de dernières nouvelles ou de décès

Modélisation de la variable latente

- Soit une variable binaire latente $Z \in \{0, 1\}$

Modélisation de la variable latente

- Soit une variable binaire latente $Z \in \{0, 1\}$
- $Z = 1$ correspond à un mauvais pronostic, $Z = 0$ à un bon pronostic

Modélisation de la variable latente

- Soit une variable binaire latente $Z \in \{0, 1\}$
- $Z = 1$ correspond à un mauvais pronostic, $Z = 0$ à un bon pronostic
- Soit la probabilité conditionnelle $\pi(x) = P[Z = 0|X = x]$ de Z sachant X d'avoir un bon pronostic en fonction des variables x

Modélisation de la variable latente

- Soit une variable binaire latente $Z \in \{0, 1\}$
- $Z = 1$ correspond à un mauvais pronostic, $Z = 0$ à un bon pronostic
- Soit la probabilité conditionnelle $\pi(x) = P[Z = 0|X = x]$ de Z sachant X d'avoir un bon pronostic en fonction des variables x
- La distribution des durées T suit le modèle de mélange suivant :
$$f(t) = \pi(x)f_0(t; \alpha_0) + (1 - \pi(x))f_1(t; \alpha_1)$$

Modélisation de la variable latente

- Soit une variable binaire latente $Z \in \{0, 1\}$
- $Z = 1$ correspond à un mauvais pronostic, $Z = 0$ à un bon pronostic
- Soit la probabilité conditionnelle $\pi(x) = P[Z = 0|X = x]$ de Z sachant X d'avoir un bon pronostic en fonction des variables x
- La distribution des durées T suit le modèle de mélange suivant :
$$f(t) = \pi(x)f_0(t; \alpha_0) + (1 - \pi(x))f_1(t; \alpha_1)$$
- avec $\pi_0(X_i) = \mathbb{P}(Z_i = 0|X_i) = \frac{1}{1 + e^{-X_i^T \beta}}$

Modélisation de la variable latente

- Soit une variable binaire latente $Z \in \{0, 1\}$
- $Z = 1$ correspond à un mauvais pronostic, $Z = 0$ à un bon pronostic
- Soit la probabilité conditionnelle $\pi(x) = P[Z = 0|X = x]$ de Z sachant X d'avoir un bon pronostic en fonction des variables x
- La distribution des durées T suit le modèle de mélange suivant :
$$f(t) = \pi(x)f_0(t; \alpha_0) + (1 - \pi(x))f_1(t; \alpha_1)$$
- avec $\pi_0(X_i) = \mathbb{P}(Z_i = 0|X_i) = \frac{1}{1 + e^{-X_i^\top \beta}}$
- Exemple : $\ell_k \sim \mathcal{G}(\alpha_k)$ de densité $\forall t \in \mathbb{N}^*, f_k(t) = \alpha_k(1 - \alpha_k)^{t-1}$

Modèle

- Paramètres : $\theta = (\alpha_0, \alpha_1, \beta)$

Modèle

- Paramètres : $\theta = (\alpha_0, \alpha_1, \beta)$
- Notations : $T^c = T \wedge C$ et $\Delta = \mathbb{1}_{\{T \leq C\}}$

Modèle

- Paramètres : $\theta = (\alpha_0, \alpha_1, \beta)$
- Notations : $T^c = T \wedge C$ et $\Delta = \mathbb{1}_{\{T \leq C\}}$
- Cohorte de n patients :

$$(x_1, t_1^c, \delta_1), \dots, (x_n, t_n^c, \delta_n) \in \mathbb{R}^d \times \mathbb{N}^* \times \{0, 1\}$$

Modèle

- Paramètres : $\theta = (\alpha_0, \alpha_1, \beta)$
- Notations : $T^c = T \wedge C$ et $\Delta = \mathbb{1}_{\{T \leq C\}}$
- Cohorte de n patients :

$$(x_1, t_1^c, \delta_1), \dots, (x_n, t_n^c, \delta_n) \in \mathbb{R}^d \times \mathbb{N}^* \times \{0, 1\}$$

- Vraisemblance du modèle :

$$\begin{aligned} \ell_n(\theta) &= \ell_n(\theta ; \mathbf{t}^c, \boldsymbol{\delta}) \\ &= -n^{-1} \sum_{i=1}^n \log \left[\{ \pi_{\beta}(x_i) f_0(t_i^c; \alpha_0) + (1 - \pi_{\beta}(x_i)) f_1(t_i^c; \alpha_1) \} \bar{G}(t_i^{c-}) \right]^{\delta_i} \\ &\quad \times \left[\{ \pi_{\beta}(x_i) \bar{F}_0(t_i^{c-}; \alpha_0) + (1 - \pi_{\beta}(x_i)) \bar{F}_1(t_i^{c-}; \alpha_1) \} g(t_i^c) \right]^{1-\delta_i} \end{aligned}$$

Inférence

- Minimisation de l'objectif :

$$\ell_n^{\text{pen}}(\theta) = \ell_n(\theta) + \gamma((1 - \eta)\|\beta\|_1 + \frac{\eta}{2}\|\beta\|_2^2)$$

Inférence

- Minimisation de l'objectif :

$$\ell_n^{\text{pen}}(\theta) = \ell_n(\theta) + \gamma((1 - \eta)\|\beta\|_1 + \frac{\eta}{2}\|\beta\|_2^2)$$

- Méthode : algorithme Quasi-Newton Expectation Maximization (QNEM) : combinaison entre un EM et un L-BFGS-B.

Inférence

- Minimisation de l'objectif :

$$\ell_n^{\text{pen}}(\theta) = \ell_n(\theta) + \gamma((1 - \eta)\|\beta\|_1 + \frac{\eta}{2}\|\beta\|_2^2)$$

- Méthode : algorithme Quasi-Newton Expectation Maximization (QNEM) : combinaison entre un EM et un L-BFGS-B.
- E-step : Calcul de $Q_n(\theta, \theta^{(l)}) = \mathbb{E}_{\theta^{(l)}}[\ell_n^{\text{comp}}(\theta) | \mathbf{t}^c, \delta]$.

Inférence

- Minimisation de l'objectif :

$$\ell_n^{\text{pen}}(\theta) = \ell_n(\theta) + \gamma((1 - \eta)\|\beta\|_1 + \frac{\eta}{2}\|\beta\|_2^2)$$

- Méthode : algorithme Quasi-Newton Expectation Maximization (QNEM) : combinaison entre un EM et un L-BFGS-B.
- E-step : Calcul de $Q_n(\theta, \theta^{(l)}) = \mathbb{E}_{\theta^{(l)}}[\ell_n^{\text{comp}}(\theta) | \mathbf{t}^c, \delta]$.
- M-step : on maximise la quantité précédente. Pour la mise à jour de β : minimisation d'un problème convexe à l'aide du L-BFGS-B.

Inférence

- Minimisation de l'objectif :

$$\ell_n^{\text{pen}}(\theta) = \ell_n(\theta) + \gamma((1 - \eta)\|\beta\|_1 + \frac{\eta}{2}\|\beta\|_2^2)$$

- Méthode : algorithme Quasi-Newton Expectation Maximization (QNEM) : combinaison entre un EM et un L-BFGS-B.
- E-step : Calcul de $Q_n(\theta, \theta^{(l)}) = \mathbb{E}_{\theta^{(l)}}[\ell_n^{\text{comp}}(\theta) | \mathbf{t}^c, \delta]$.
- M-step : on maximise la quantité précédente. Pour la mise à jour de β : minimisation d'un problème convexe à l'aide du L-BFGS-B.
- Propriété de convergence de l'algorithme vérifiée.

Principes

- Simulations de Monte Carlo

Principes

- Simulations de Monte Carlo
- Génération des covariables

Principes

- Simulations de Monte Carlo
- Génération des covariables
- Génération des temps suivant 3 modèles : modèle de mélange, modèle de Cox, modèle CURE
- Comparaison de la performance des différentes méthodes d'inférence des paramètres : QNEM, Cox pénalisé et CURE pénalisé

Génération des covariables

- Soit le vecteur de coefficient suivant :
 $\beta = (\underbrace{K, \dots, K}_s, 0, \dots, 0) \in \mathbb{R}^d$, avec $K \in \mathbb{R}$ la valeur des coefficients actifs, $s \in \{1, \dots, d\}$

Génération des covariables

- Soit le vecteur de coefficient suivant :
 $\beta = (\underbrace{K, \dots, K}_s, 0, \dots, 0) \in \mathbb{R}^d$, avec $K \in \mathbb{R}$ la valeur des coefficients actifs, $s \in \{1, \dots, d\}$
- Soit $\pi_0 \in [0, 1]$ la proportion des individus à risque faible.

Génération des covariables

- Soit le vecteur de coefficient suivant :

$$\beta = (\underbrace{K, \dots, K}_s, 0, \dots, 0) \in \mathbb{R}^d$$
, avec $K \in \mathbb{R}$ la valeur des coefficients actifs, $s \in \{1, \dots, d\}$
- Soit $\pi_0 \in [0, 1]$ la proportion des individus à risque faible.
- Les covariables sont générées suivant un modèle de mélange gaussien : $x_{ij} \leftarrow x_{ij} \pm \text{gap}$ for $j \in \{1, \dots, s, \dots, \lfloor (d-s)r_{cf} \rfloor\}$

Génération des covariables

- Soit le vecteur de coefficient suivant :
 $\beta = (\underbrace{K, \dots, K}_s, 0, \dots, 0) \in \mathbb{R}^d$, avec $K \in \mathbb{R}$ la valeur des coefficients actifs, $s \in \{1, \dots, d\}$
- Soit $\pi_0 \in [0, 1]$ la proportion des individus à risque faible.
- Les covariables sont générées suivant un modèle de mélange gaussien : $x_{ij} \leftarrow x_{ij} \pm \text{gap}$ for $j \in \{1, \dots, s, \dots, \lfloor (d-s)r_{cf} \rfloor\}$
- Puis on génère $Z_i \sim \mathcal{B}(1 - \pi_\beta(x_i))$

Génération des covariables

- Soit le vecteur de coefficient suivant :
 $\beta = (\underbrace{K, \dots, K}_s, 0, \dots, 0) \in \mathbb{R}^d$, avec $K \in \mathbb{R}$ la valeur des coefficients actifs, $s \in \{1, \dots, d\}$
- Soit $\pi_0 \in [0, 1]$ la proportion des individus à risque faible.
- Les covariables sont générées suivant un modèle de mélange gaussien : $x_{ij} \leftarrow x_{ij} \pm \text{gap}$ for $j \in \{1, \dots, s, \dots, \lfloor (d-s)r_{cf} \rfloor\}$
- Puis on génère $Z_i \sim \mathcal{B}(1 - \pi_\beta(x_i))$
- Et enfin $T_i \sim \mathcal{G}(\alpha_{Z_i})$

Génération des temps et inférence des paramètres : le modèle de Cox

- Cox PH modèle :

$$\ell_n^{\text{COX}}(\beta) = -n^{-1} \sum_{i=1}^n \left(\mathbf{x}_i^{\top} \beta - \log \sum_{i': t_{i'}^c \geq t_i^c} \exp(\mathbf{x}_{i'}^{\top} \beta) \right)^{\delta_i}$$

Génération des temps et inférence des paramètres : le modèle de Cox

- Cox PH modèle :

$$\ell_n^{\text{COX}}(\beta) = -n^{-1} \sum_{i=1}^n \left(\mathbf{x}_i^\top \beta - \log \sum_{i': t_{i'}^c \geq t_i^c} \exp(\mathbf{x}_{i'}^\top \beta) \right)^{\delta_i}$$

- On utilise `glmnet` pour minimiser

$$\ell_n^{\text{COX}}(\beta) + \gamma \left((1 - \eta) \|\beta\|_1 + \frac{\eta}{2} \|\beta\|_2^2 \right)$$

Génération des temps et inférence des paramètres : le modèle de Cure

- Modèle de mélange qui suppose qu'une fraction de la population est guérie ($T = \infty$).

Génération des temps et inférence des paramètres : le modèle de Cure

- Modèle de mélange qui suppose qu'une fraction de la population est guérie ($T = \infty$).
- Cas particulier de notre modèle, aucune implémentation avec régularisation connue donc on a utilisé notre modèle avec $\alpha_0 = 0$ pour des lois géométriques.

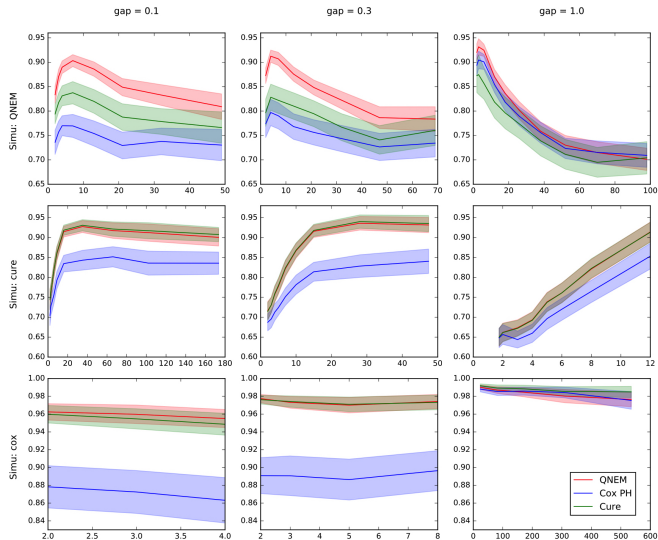
Métrique utilisée pour l'estimation de la performance

- AUC(t)

Métrique utilisée pour l'estimation de la performance

- AUC(t)
- C-index $\mathcal{C}_\tau = \mathbb{P}[M_i > M_j | T_i^c < T_j^c, T_i^c < \tau]$

Résultats des simulations



Identification des variables à l'origine de réhospitalisation précoce des patients atteints de crise vaso-occlusive

- Drépanocytose ou "anémie falciforme" : maladie génétique rare touchant le globule rouge

Identification des variables à l'origine de réhospitalisation précoce des patients atteints de crise vaso-occlusive

- Drépanocytose ou "anémie falciforme" : maladie génétique rare touchant le globule rouge
- Crise vaso-occlusive : complication aiguë de la drépanocytose, à l'origine de douleurs intenses

Identification des variables à l'origine de réhospitalisation précoce des patients atteints de crise vaso-occlusive

- Drépanocytose ou "anémie falciforme" : maladie génétique rare touchant le globule rouge
- Crise vaso-occlusive : complication aiguë de la drépanocytose, à l'origine de douleurs intenses
- HEGP = centre de référence du traitement des crises vaso-occlusives

Identification des variables à l'origine de réhospitalisation précoce des patients atteints de crise vaso-occlusive

- Drépanocytose ou "anémie falciforme" : maladie génétique rare touchant le globule rouge
- Crise vaso-occlusive : complication aiguë de la drépanocytose, à l'origine de douleurs intenses
- HEGP = centre de référence du traitement des crises vaso-occlusives
- Rechute fréquente à la sortie de l'hospitalisation

Identification des variables à l'origine de réhospitalisation précoce des patients atteints de crise vaso-occlusive

- Drépanocytose ou "anémie falciforme" : maladie génétique rare touchant le globule rouge
- Crise vaso-occlusive : complication aiguë de la drépanocytose, à l'origine de douleurs intenses
- HEGP = centre de référence du traitement des crises vaso-occlusives
- Rechute fréquente à la sortie de l'hospitalisation
- Litterature pauvre et contradictoire sur les facteurs prédictifs à l'origine des réhospitalisations

Données cliniques disponibles

- Réutilisation des données de soin

Données cliniques disponibles

- Réutilisation des données de soin
- HEGP = hôpital informatisé avec un entrepôt de données cliniques

Données cliniques disponibles

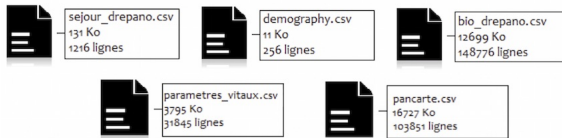
- Réutilisation des données de soin
- HEGP = hôpital informatisé avec un entrepôt de données cliniques
- Pour garantir une homogénéité des pratiques médicales, extraction des données de 2008 à 2015

Données cliniques disponibles

- Réutilisation des données de soin
- HEGP = hôpital informatisé avec un entrepôt de données cliniques
- Pour garantir une homogénéité des pratiques médicales, extraction des données de 2008 à 2015
- Extraction limitée aux sources suivantes : biologie, paramètres vitaux, dates des séjours, âge et sexe.

Organisation des données

- Données hétérogènes, différentes sources

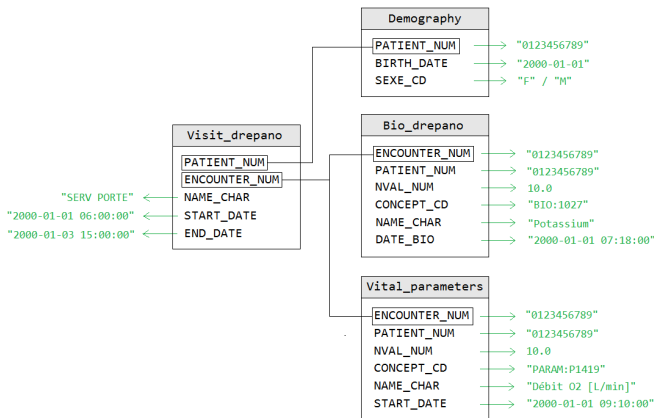


Organisation des données

- Données hétérogènes, différentes sources
- 1 séjour / patient : "choix aléatoire par classe", taux de rechute de 12,84%

Organisation des données

● Réorganisation des données



Organisation des données

● Création d'un fichier JSON

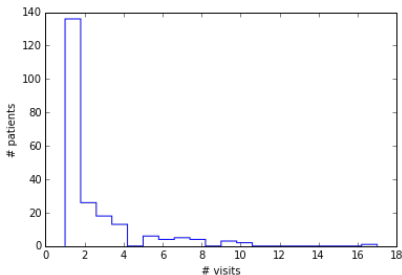
```

Patients = {
  '1' = {
    'patient_num' = '0123456789',
    'sex' = 'F/M',
    'birth_date' = '2000-01-01',
    'visits' = {
      '1' = {
        'encounter_num' = '0123456789',
        'age' = '25',
        'rea' = '0/1',
        'duration' = '13',
        'previous_visit' = 'none/10',
        'next_visit' = 'none/15',
        'details' = {
          '1' = {
            'service' = 'SERV PORTE',
            'start_date' = '2000-01-01 05:00:00',
            'end_date' = '2000-05-01 10:00:00'
          },
          '2' = {...}
        },
        'bio-infos' = {
          '1' = {
            'name_char' = 'Potassium',
            'val' = {
              '1' = {
                'nval_num' = '10.0',
                'concept_cd' = '810:1027',
                'date_bio' = '2000-01-01 11:00:00'
              },
              '2' = {...}
            }
          },
          '2' = {...}
        },
        'vital_parameters' = {
          '1' = {
            'name_char' = 'Débit O2 [L/min]',
            'val' = {
              '1' = {
                'nval_num' = '10.0',
                'concept_cd' = 'PARAM:P1419',
                'date_bio' = '2000-01-01 10:30:00'
              },
              '2' = {...}
            }
          },
          '2' = {...}
        }
      },
      '2' = {...}
    }
  },
  '2' = {...}
}

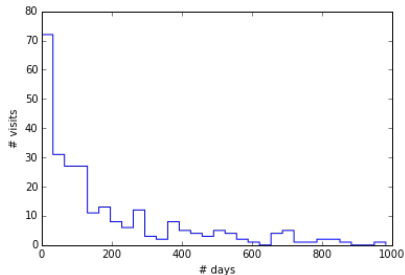
```


Analyse descriptive

● 218 patients, 479 visites



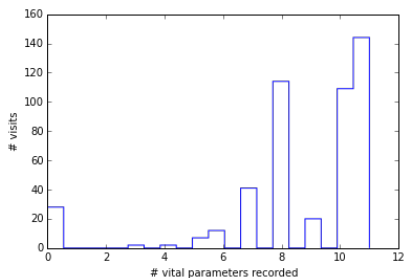
Nombre de visites par patient



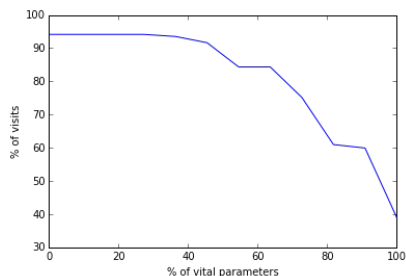
Temps inter-visites

Les paramètres vitaux

● Les paramètres vitaux



Nombre de points
par paramètre vital



Variables biologiques en commun
à tous les séjours

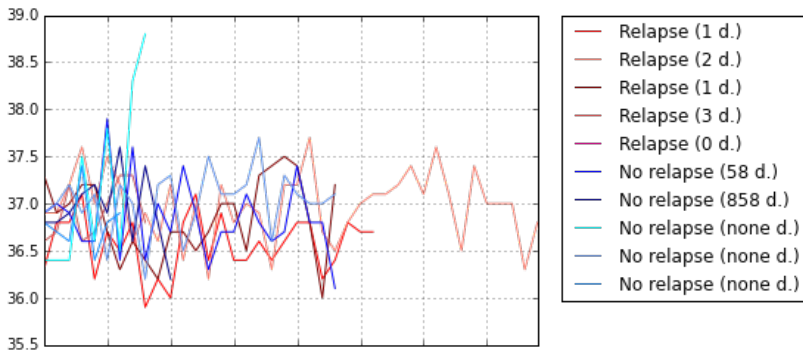
Les paramètres vitaux

- Les paramètres vitaux
- 9 variables gardées

Pression artérielle systolique [mmHg], notée PA max ;
pression artérielle diastolique [mmHg], notée PA min ;
saturation en oxygène [%] ; Douleur EVA [U] ; fréquence
respiratoire [mvt/min] ; température [°C] ; poids [kg] ;
oxygène [L/min] ; fréquence cardiaque [bpm]

Les paramètres vitaux

- Les paramètres vitaux
- 9 variables gardées
- Problème : durées $\neq$ et alignement



Evolution de la température au cours de 10 séjours

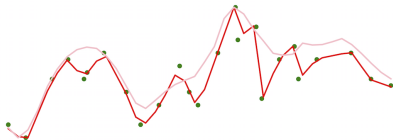
Les paramètres vitaux

- Rescaling entre $[0,1]$

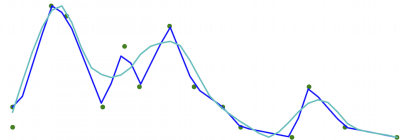
$$\forall j \in \llbracket 1, N \rrbracket, x_j = \frac{x_j - \min_k(x_k)}{\max_k(x_k) - \min_k(x_k)}$$

Les paramètres vitaux

- Rescaling entre $[0,1]$
- Interpolation : p valeurs
Puis filtre de Savitzky–Golay : p' valeurs de gradient

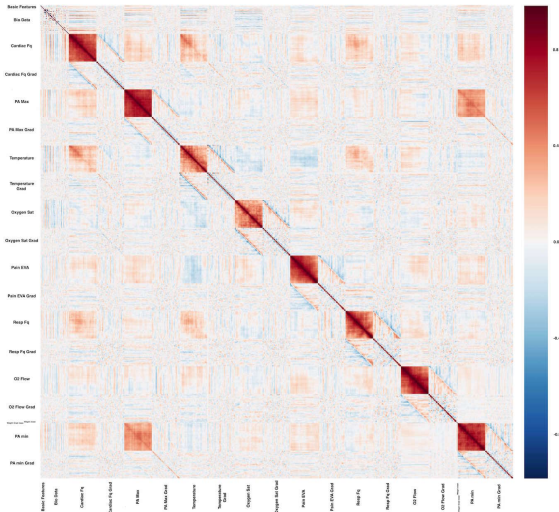


Fréquence respiratoire pour
un patient positif



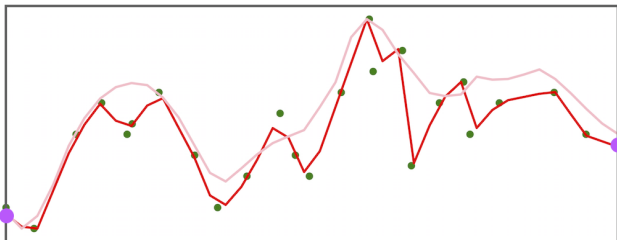
Fréquence respiratoire pour
un patient négatif

Preprocessing



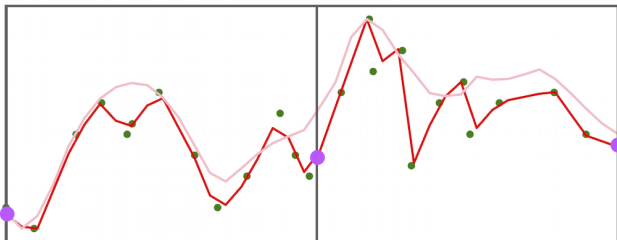
Sélection de p et p'

Procédure suivie :



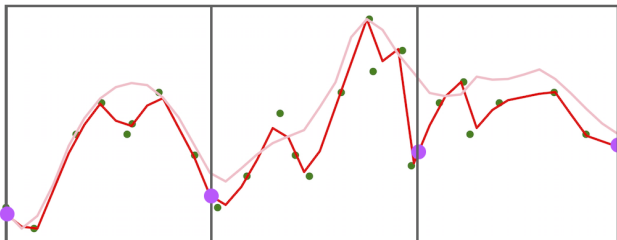
Sélection de p et p'

Procédure suivie :



Sélection de p et p'

Procédure suivie :



Sélection de p et p'

● Résultats après 50 itérations

Concept	Nombre de points	Ecart type
Fq cardiaque	28	0,3
▽ Fq cardiaque	26	0
PA max	28	0
▽ PA max	26	0
Température	30	8,4
▽ Température	18	0
Saturation O_2	34	4,8
▽ Saturation O_2	26	0
Douleur EVA	21	0,7
▽ Douleur EVA	26	0
Débit O_2	16	4,6
▽ Débit O_2	26	0
Fq respiratoire	21	5,3
▽ Fq respiratoire	26	0
PA min	28	0
▽ PA min	26	0

Sélection de p et p'

- Résultats après 50 itérations
- Dimension de l'espace : 446

Sélection de p et p'

- Résultats après 50 itérations
- Dimension de l'espace : 446
- Données manquantes remplacées par les moyennes

Sélection des variables par procédure de Benjamini-Hochberg

- Tests univariés : 446 *p-values*

Sélection des variables par procédure de Benjamini-Hochberg

- Tests univariés : 446 *p-values*
- Sélection des variables en fonction du FDR

On conserve k features avec

$$k = \operatorname{argmax} \left\{ j : p(j) \leq \frac{\alpha j}{K} \right\}$$

Sélection des variables par procédure de Benjamini-Hochberg

- Tests univariés : 446 *p-values*
- Sélection des variables en fonction du FDR
- $\alpha = 0,9$; 50 itérations

29/36

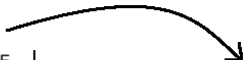
Binarisation

- Binarisation par quartiles ; $dim = 156 \sim 4 \times 45$

...	F_j	...
	:	
	:	
	12,5	
	9,2	
	3,1	
	8,7	
	:	
	:	

Binarisation

- Binarisation par quartiles ; $dim = 156 \sim 4 \times 45$

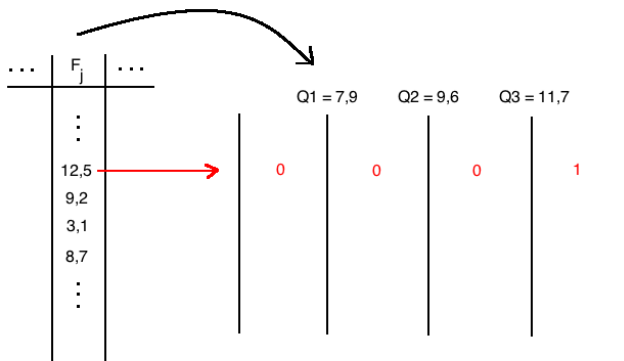


...	F_j	...
	$\vdots$	
	12,5	
	9,2	
	3,1	
	8,7	
	$\vdots$	

Q1 = 7,9 Q2 = 9,6 Q3 = 11,7

Binarisation

- Binarisation par quartiles ; $dim = 156 \sim 4 \times 45$

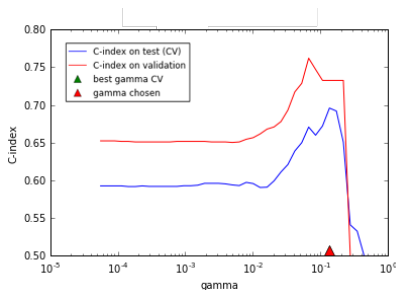


Binarisation

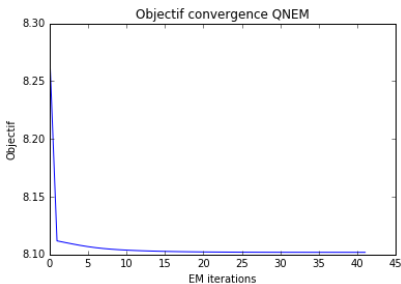
- Binarisation par quartiles ; $dim = 156 \sim 4 \times 45$

...	F_j	...		$Q1 = 7,9$	$Q2 = 9,6$	$Q3 = 11,7$	
	$\vdots$			$\vdots$	$\vdots$	$\vdots$	$\vdots$
	12,5	$\rightarrow$	0	0	0	1	
	9,2	$\rightarrow$	0	1	0	0	
	3,1	$\rightarrow$	1	0	0	0	
	8,7	$\rightarrow$	0	1	0	0	
	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Résultats sur RedCVO



(a) Learning curves



(b) Convergence

Résultats sur RedCVO

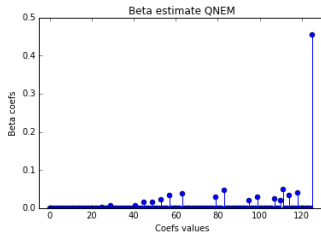


FIGURE: Beta qnem

Features	Effets
previous visit (686, 2412]	51.7 %
Frquence respiratoire [<i>mvt/min</i>] gradient 11 (0, 0.0695]	5.6 %
Saturation en oxygène gradient 13 (−0.089, 0.0345]	5.2 %
Dbit O2 [<i>L/min</i>] gradient 25 [−1.659, −0.0307]	4.5 %
Température [C] point 26 (36.724, 36.93]	4.3 %
Température [C] point 19 (36.767, 36.922]	3.9 %
Dbit O2 [<i>L/min</i>] gradient 2 [−0.623, −0.0212]	3.7 %
Douleur EVA point 17 (0.542, 2.223]	3.2 %

Résultats sur RedCVO

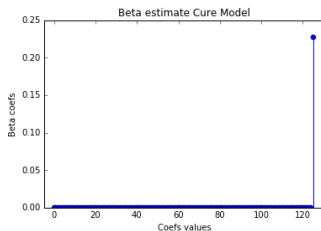


FIGURE: Beta Cure

Features	Effets
previous visit (686, 2412]	100 %

Résultats sur RedCVO

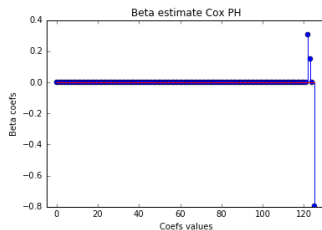


FIGURE: Beta Cox PH

Features	Effets
previous visit (686, 2412]	63.3 %
previous visit [1, 135]	24.5 %
previous visit (135, 364]	12.2 %

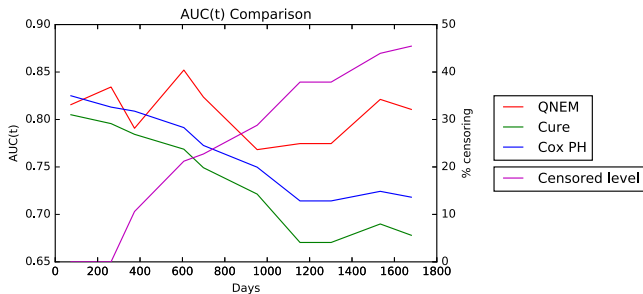


FIGURE: AUC(t)

Models	C-index
QNEM	0.77
Cure	0.73
Cox PH	0.75

Conclusion

- Originalité du modèle : modèle de mélange pénalisé pour données censurées

Conclusion

- Originalité du modèle : modèle de mélange pénalisé pour données censurées
- Très bonnes performances par rapport au modèle de Cox et au modèle de CURE

Conclusion

- Originalité du modèle : modèle de mélange pénalisé pour données censurées
- Très bonnes performances par rapport au modèle de Cox et au modèle de CURE
- Prochaine étape : Extension à d'un mélange à K composantes, avec optimisation de K .