

UV ELECTIVE 501 : Data Science

Présentation du Défi 3: Zombie apocalypse

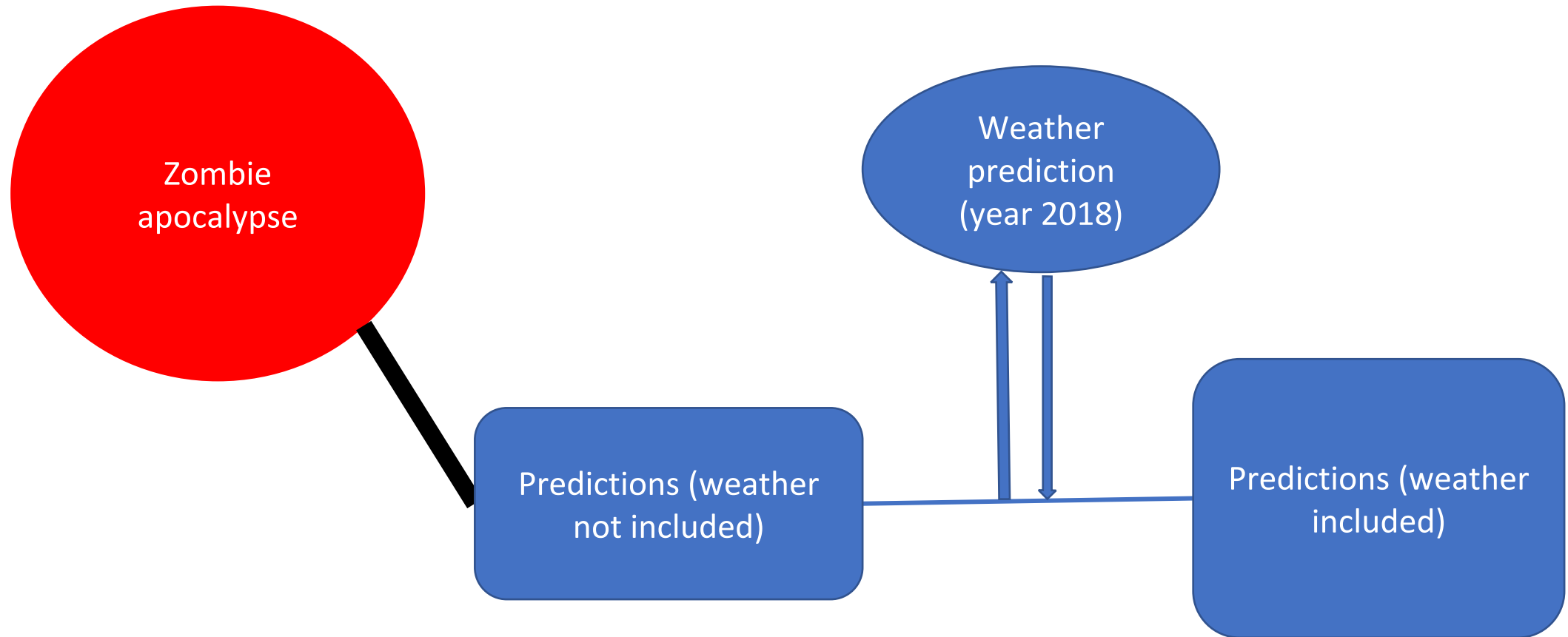
Auteurs:

- **Badr Labrihli**
- **Marwane Mesbahi**
- **Omar Al banchaabouchi**

Sommaire :

- Introduction
- Méthodologie
- 1^{ère} étape : résultats sans l'introduction de la température
- Prédiction de la température
- 2^{ème} étape : résultats (température incluse)
- Conclusion
- Bibliographie

Introduction



Méthodologie :

Usage de la méthode CRISP-DM :

- I. Prédiction de l'éventuelle atteinte des zombies sur Brest
- II. Prédiction de la température

I – Prédiction de l'éventuelle atteinte des zombies sur Brest :

Business Understanding : Zombie apocalypse, predict whether zombies are going to reach Brest or not.

Data Understanding :

- a BMP file
- Information about zombies (speed..)

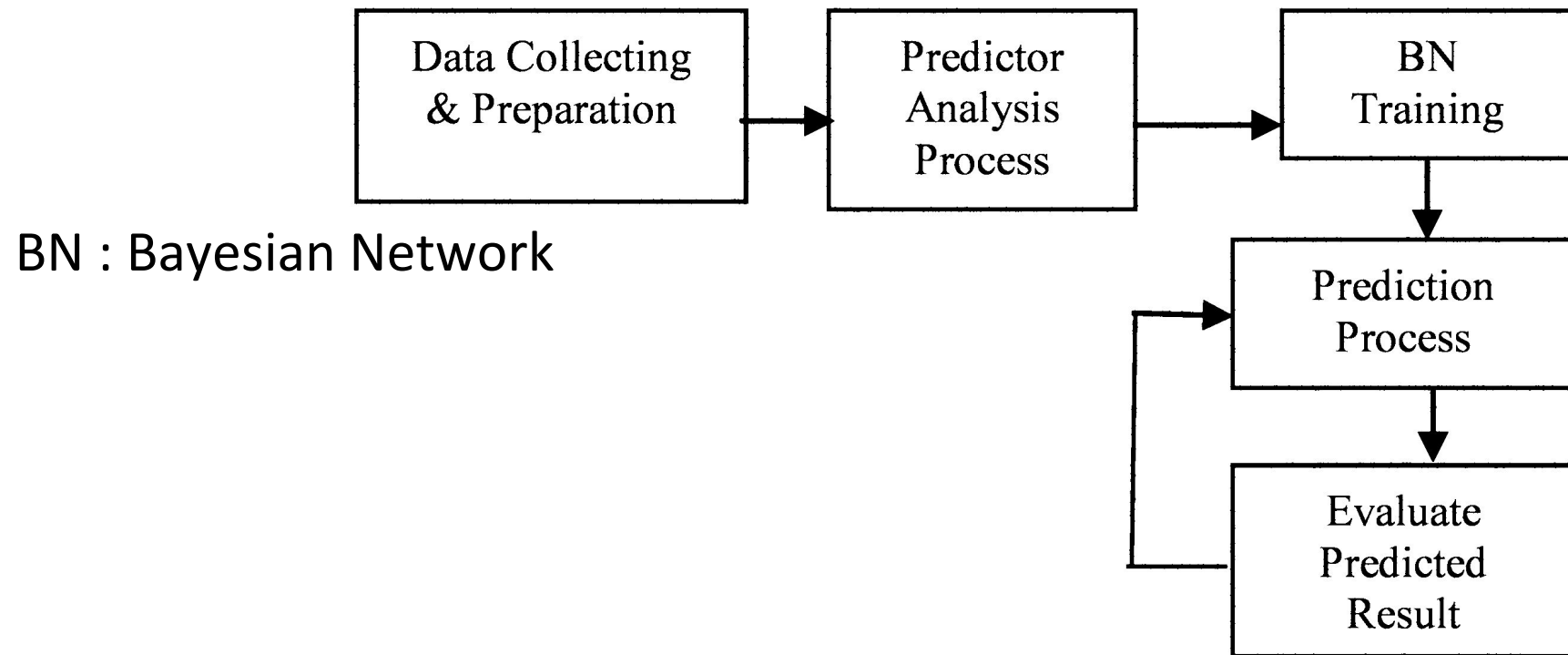
Data Preparation : Creating a weighted graph

Modeling : Shortest path

Evaluation

Deployment

II –Prédiction de la température :



Prédictions (sans température)

I- Première approche (difficile) :

- Modéliser l'extension de l'épidémie (Zombie) à l'aide des chaînes de markov:

$$S' = -\beta SZ$$

$$Z' = +\beta SZ + \zeta R - \alpha SZ$$

$$R' = +\alpha SZ - \zeta R$$

- Se baser sur la formule d'Euler simplifiée pour représenter le passage d'un état à un autre:

$$u'(t) = \lim_{\Delta t \rightarrow 0} \frac{u(t + \Delta t) - u(t)}{\Delta t}$$

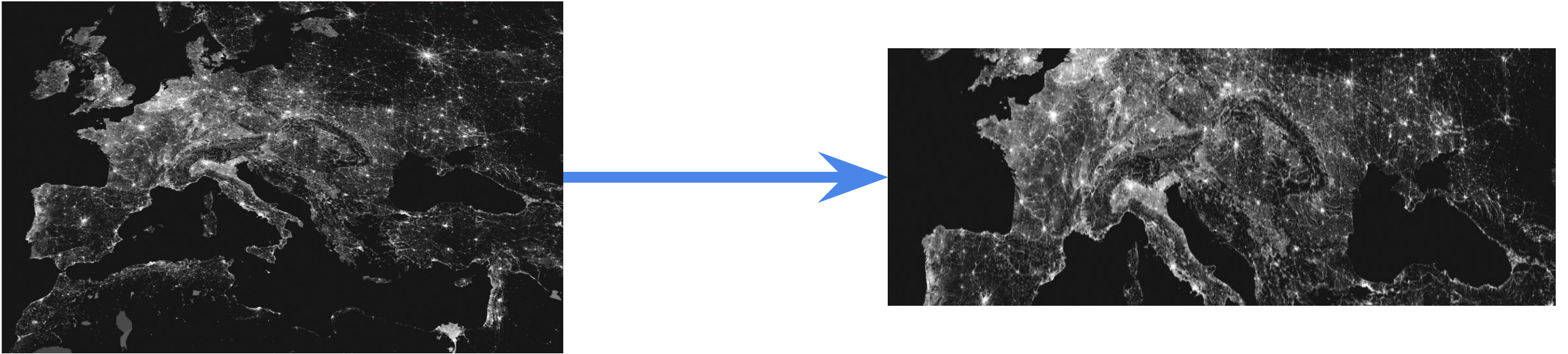
$$u' \Delta t + u(t) = u(t + \Delta t)$$

$$f(u) = u'(t)$$

$$u_{n+1} = f(u) \Delta t + u_n$$

Prédictions (sans température)

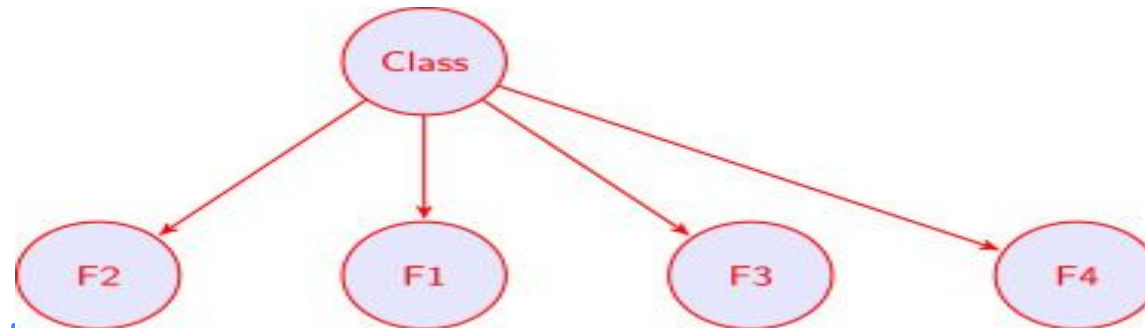
I- Deuxième approche (simplifiée) :



- Réduction de la taille de l'image traitée.
- Génération d'un graphe pondéré à partir de l'image.
- Déduction du plus court chemin par l'algorithme A*.
- Vérification du cas où un zombie meurt s'il prend ce chemin.

Prédiction de la température :

- Création d'un arbre dont les feuilles sont les éléments à prédire, et les racines sont les éléments du Training set. Exemple:



- Méthode de prédiction probabiliste basée sur la formule de Bayes:

$$P(c | x) = \frac{P(x | c) P(c)}{P(x)}$$

Labels for the formula:

- $P(c | x)$ is labeled **Posterior Probability**.
- $P(x | c)$ is labeled **Likelihood**.
- $P(c)$ is labeled **Class Prior Probability**.
- $P(x)$ is labeled **Predictor Prior Probability**.

$$P(c | X) = P(x_1 | c) \times P(x_2 | c) \times \dots \times P(x_n | c) \times P(c)$$

Prédictions (température incluse) :

Modification de la fonction vitesse zombie :

- **Vitesse = Vitesse_été*Température/18** si $0 \leq \text{Température} \leq 18$
- **Vitesse = Vitesse_été** si $18 \leq \text{Température}$
- **Vitesse = 0** sinon

Problèmes rencontrés:

- L'image bmp fournie est de très grande taille.
- Les données météorologiques sont difficiles à lire et à traiter.
- La session de l'école ne nous permet pas de stocker assez de données pour le traitement.

Résultats

Notre approche nous a permis de conclure que la ville de Brest sera infectée après 676 jours.

Bibliographie:

- <https://www.sciencedirect.com/science/article/pii/S0952197617301513>
- <https://arxiv.org/pdf/1311.6376.pdf>
- <http://maxberggren.se/2014/11/27/model-of-a-zombie-outbreak/>
- http://www.saedsayad.com/naive_bayesian.htm