



**Laboratoire de Mathématiques et Informatique pour la Complexité et les Systèmes  
MICS**

**Présente**

## **L'AVIS DE SOUTENANCE**

**De Madame Elvire ROBLIN**

Laboratoire MICS, CentraleSupélec, Université Paris Saclay, soutiendra publiquement  
ses travaux de thèse de doctorat intitulés :

**"Survival Prediction using Artificial Neural Networks on Censored Data"**

**Le 21 mars 2023 à 08h30**

À l'Institut Gustave Roussy, dans la **salle Diamant 62** - Bâtiment de Médecine  
Moléculaire en présentiel et en distanciel (pour le présentiel merci de bien vouloir  
contacter en avance l'assistante du Laboratoire à l'adresse mail :  
Fabienne.brosse@centralesupelec.fr)

*Si vous souhaitez assister à la soutenance en distanciel, veuillez contacter en avance l'assistante du laboratoire pour obtenir  
le lien.*

### **Membres du jury :**

Hugues Talbot, Professeur des universités (HDR), CVN, CentraleSupélec, Université Paris Saclay

Anne-Laure Boulesteix, Professeur des universités (HDR), Medizinische Fakultät – Ludwig Maximilians  
Universität München (Rapporteur)

Pascal Roy, Professeur des universités – praticien hospitalier (HDR), Université Claude Bernard Lyon 1  
(Examineur)

Axel Benner, Directeur de recherche, Deutsches Krebsforschungszentrum (Examineur)

Angelina Roche, Maître de conférences, CEREMADE, Université Paris Dauphine (Examinatrice)

Paul-Henry Cournède, Professeur (HDR), MICS, CentraleSupélec (Directeur de thèse)

Stefan Michiels, Directeur de recherche (HDR), Oncostat, Institut Gustave Roussy (Co-directeur de  
thèse)

## **Résumé :**

En oncologie, l'analyse des données de survie a des applications multiples : analyse de la survie d'un groupe de patients, comparaison des effets de différents traitements, etc. Les données de survie présentent la caractéristique de ne pas être observées complètement, l'événement d'intérêt n'étant pas nécessairement survenu pour tous les patients à la fin de la période d'observation : les données sont dites censurées. Dans les modèles pronostiques, on utilise traditionnellement un modèle à risques proportionnels de Cox pour analyser les données de survie. Ce modèle repose sur des hypothèses restrictives qui ne permettent pas de prendre en compte facilement l'analyse d'un nombre toujours croissant de facteurs pronostiques et prédictifs, ni d'inclure toutes leurs interactions ou leurs effets non linéaires. De nouveaux modèles s'appuyant sur les réseaux de neurones artificiels peuvent être employés pour faire face à ces limites. L'objectif de cette thèse est d'utiliser ces modèles et de les adapter au contexte des données censurées. Dans un premier temps, nous implémentons des modèles de réseaux de neurones permettant d'effectuer des prédictions de survie. Nous comparons des méthodes existantes à nos méthodes proposées et construites à l'aide de pseudo-observations. Dans un second temps, nous établissons une mesure de l'incertitude des prédictions réalisées avec ces réseaux de neurones. Elles prennent la forme d'intervalles de confiance. Enfin, nous mettons en place des recommandations de traitement individualisées.

## **Abstract :**

In oncology, analyzing survival data is of primary importance in epidemiological studies and clinical research. The most commonly used approach to deal with the censored observations present in this type of study is the Cox proportional hazards model. The model is based on particular assumptions which may be hard to assume for a high number of candidate covariates to evaluate (proportional hazards, linearity of effects, interactions between covariates). Recently a new family of models, Artificial Neural Networks, has been proposed to model complex patterns and prediction problems, and may be used in the context of survival prediction. They have been widely used in different areas of science. This Ph.D. aims to provide new methods of survival analysis using neural networks. The objective of the first part of the Ph.D. project is to compare existing neural network models adapted to survival with our proposed method based on pseudo observations as a solution to the problem of censored data. The second part focuses on measuring the uncertainty of our predictions using confidence intervals. Finally, a treatment recommendation rule based on a neural network is set up in order to provide treatment recommendations based on the patient's features.