

## Abstract

Banks are key players in our economies. They enable the financing of projects requiring large amounts of capital. Failures create a domino effect in the banking system, which can lead to a major financial panic. The historical financial statements that allow us to predict in advance their capacity to face possible crisis shocks thanks to the variables they contain.

Our work firstly consists of a set of statistical calculations allowing us to verify whether there is a significant difference between the banks that have failed and those that have survived, based on their capital and liquidity. In this case, we used a t-test and a z-test. Secondly, we used two machine learning models: logistic regression and neural network. Both receive capital and cash of all banks as input values. The result at the end will be a binary value that will tell us if the bank has gone bankrupt or survived.

After implementing our two models, we obtained an accuracy of 69% for the logistic regression, while the neural network performed less efficiently because of its high number of false positives.

This led us to conclude that we would need more data from banks and even macroeconomic variables to make more accurate predictions. Deep learning models could also be used to improve them.

## Abstract

Les banques sont des acteurs clés de nos économies. Elles permettent le financement de projets nécessitant des capitaux importants. Les défaillances créent un effet domino dans le système bancaire, pouvant mener à une importante panique financière. Les états financiers historiques qui permet de prévoir en amont leur capacité d'affronter d'éventuels chocs de crise grâce aux variables qu'ils comportent.

Notre travail consiste, en premier lieu, en un ensemble de calculs statistiques permettant de vérifier s'il y a une différence significative entre les banques qui ont fait faillite et celles qui ont survécu en nous basant sur leurs capitaux et leurs liquidités. En l'occurrence, il s'agit du t-test et du z-test. En second lieu, nous avons utilisé deux modèles de *machine learning*, la régression logistique et le réseau neuronal. Tous les deux reçoivent en valeur d'entrée le capital et la liquidité. Le résultat consistera en une valeur binaire qui nous dira si la banque a fait faillite ou survécu.

Après l'implémentation de nos deux modèles, nous avons obtenu, pour la régression logistique, une précision de 69 %, tandis que le réseau de neurones s'est avéré moins performant en raison de son nombre élevé de faux positifs.

Ceci nous a permis de conclure que nous aurions besoin de plus de données bancaires, voire macroéconomiques, pour réaliser une prédiction plus précise. Des modèles de *deep learning* seraient également envisageables dans l'amélioration de cette prédiction.