



Rapport de projet
Mineure recherche

**Prévision des défaillances bancaires à travers les
états financiers**

Laboratoire de recherches de l'ECE Paris-Lyon

Laboratoire de recherche de l'EBS

Référents projet : **M. Jae Yun JUN KIM**

M. Angelo RIVA

Réalisé par : **Karim ER-RACHDI**

Jorge BELMONTE-JIMENEZ

Amr SAARIG



SOMMAIRE

| | |
|---|----|
| <u>INTRODUCTION</u> | 3 |
| <u>DEVELOPPEMENT</u> | 4 |
| I. <u>Présentation des membres de l'équipe</u> | 4 |
| II. <u>Présentation du laboratoire de l'ECE</u> | 5 |
| <u>B) Histoire</u> | 5 |
| <u>A) Partenariats</u> | 6 |
| III. <u>Contexte du travail de recherche</u> | 6 |
| IV. <u>Etat de l'art</u> | 7 |
| V. <u>Nouveautés de recherche</u> | 11 |
| VI. <u>Nos contributions</u> | 13 |
| VII. <u>Partie technique et developpement</u> | 14 |
| <u>CONCLUSION</u> | 20 |
| <u>PERSPECTIVES</u> | 20 |
| <u>REFERENCES</u> | 21 |
| <u>ANNEXES</u> | 21 |

INTRODUCTION

Les banques sont des acteurs clés de nos économies. En effet, elles permettent le financement de projets nécessitant d'importants montants de capitaux (prêts immobiliers, financement d'entreprises, entre autres) et servent d'intermédiaire des politiques monétaires des banques centrales. Les nombreux changements de nos macroéconomies ont fortement modifié le fonctionnement du secteur bancaire, créant ainsi des crises bancaires et des régulations pour mieux encadrer le secteur du fait de son rôle prépondérant dans l'économie réelle.

Un défaut bancaire survient, dans 90 % des cas, quand une ou plusieurs banques ne peuvent plus satisfaire leurs obligations financières, c'est-à-dire qu'elles ne peuvent plus rémunérer leurs créanciers. Dans ce scénario, un effet contagion peut se propager, créant ainsi une crise bancaire et, dans les cas les plus graves, une récession économique. Lors d'un défaut bancaire, nous cherchons toujours à savoir ce qui a mené à l'effondrement d'une ou plusieurs banques, mais les chercheurs en économie et en finance ne parviennent pas à donner une réponse unifiée et précise.

Notre travail consistera en l'utilisation des états financiers des banques pour prévoir les défauts bancaires. En effet, les états financiers reflètent la santé financière de la banque, ses revenus, ses dépenses ainsi que ses actifs et passifs. Ceci nous permettrait, en fonction du contexte macroéconomique et d'autres variables qui seront précisées dans ce rapport, d'évaluer le risque de défaillance ou de survie d'une banque en cas de crise.

Et puisque nous utiliserons les états financiers, nous sommes contraints de répondre à la question suivante : **Comment, à partir des états financiers, allons-nous prévoir la probabilité qu'une banque fasse faillite ?**

Dans un premier temps, nous commencerons par présenter notre sujet dans le détail avec les mots clés qui le définissent, les membres de l'équipe, le laboratoire de recherche et la problématique de notre travail. Puis nous ferons une présentation de l'état de l'art qui se compose des mesures de performance et des variables de contrôle et d'état. Ensuite, nous présenterons les contributions des chercheurs les plus importantes que nous avons extraites de nos lectures. Après,

nous aborderons la méthodologie appliquée dans le cadre de notre projet. Enfin, nous terminerons ce rapport par une conclusion ainsi qu'un ensemble de perspectives de ce projet.

DEVELOPPEMENT

I. Présentation des membres de l'équipe

L'équipe de ce projet est composée de 3 étudiants de l'ECE Paris-Lyon en quatrième année et tous issus de la majeure finance quantitative.

Nous présenterons le parcours de chacun d'entre eux :

- **Karim ER-RACHDI** : à la suite de 2 années de classes préparatoires MPSI-MP au Maroc, il a décidé d'intégrer l'ECE Paris-Lyon à la rentrée 2020-2021. Il a décidé de s'orienter vers la majeure finance quantitative en raison de son intérêt prononcé par les mathématiques appliquées aux marchés financiers. Grâce à l'enseignement qu'il a reçu en classe préparatoire, il dispose d'une très grande capacité d'abstraction, d'un fort raisonnement logique et de solides compétences en calcul, probabilités et statistiques. Son enseignement à l'ECE lui a donné de bonnes bases dans certains langages tels que C, Java et Python grâce aux enseignements de programmation, intelligence artificielle et Machine-Learning. Il souhaiterait travailler dans une banque d'investissement à la suite de ses études à l'ECE Paris-Lyon.
- **Jorge BELMONTE-JIMENEZ** : il s'est orienté à l'ECE Paris-Lyon à l'issue de l'obtention de son baccalauréat. Sa scolarité à l'ECE lui a donné de fortes compétences dans beaucoup de langages de programmation tels que C, C++, Java et Python. Il a également un goût prononcé pour les concepts mathématiques et savait qu'il voulait choisir la majeure finance quantitative dû à sa passion à la fois pour les mathématiques et les marchés



financiers. Pour cette raison, il s'est orienté vers la majeure finance quantitative à son entrée en quatrième année. A l'issue de ses études, il voudrait travailler dans une banque ou une société d'investissement.

- **Amr SAARIG** : à la suite de 2 années de classes préparatoires MPSI-MP au Maroc, il a décidé d'intégrer l'ECE Paris-Lyon à la rentrée 2020-2021. Il a décidé de s'orienter vers la majeure finance quantitative en raison de son intérêt prononcé par les mathématiques appliquées aux marchés financiers. Grâce à l'enseignement qu'il a reçu en classe préparatoire, il dispose d'une très grande capacité d'analyse et de solides compétences en calcul de probabilités. Son enseignement à l'ECE lui a donné les langages de programmation tels que C, Java et Python. Il souhaiterait travailler dans un fonds ou des banques d'investissement à la suite de ses études à l'ECE Paris-Lyon.

II. Présentation du laboratoire de recherche de l'ECE Paris-Lyon

Le laboratoire de recherche de l'ECE Paris-Lyon est un laboratoire de recherche appliquée au service de la société de demain, des entreprises, des organisations et adaptée au choix du positionnement des écoles du groupe OMNES.

Le laboratoire de l'école a 3 missions principales. La première est celle de la formation étant donné qu'il se situe dans un établissement d'enseignement supérieur et fait partie intégrante de la formation d'ingénieurs en initiant les élèves de la mineure recherche à la recherche scientifique. En outre, de nombreux enseignants-chercheurs y travaillent sur plusieurs thématiques que l'on détaillera plus tard. La deuxième mission est celle de l'innovation dans laquelle les sujets de recherche s'inscrivent et prennent en compte des dimensions scientifiques, économiques, sociétales ou encore environnementales. La troisième est celle des liens avec le milieu industriel car les nouvelles innovations scientifiques qui font l'objet d'études au sein du laboratoire contribuent au développement de plusieurs industries (automobile, informatique, électronique, entre autres) qui doivent rester compétitives



et deviennent plus exigeantes face aux actuels changements technologiques.

A) Histoire

Depuis sa création, le laboratoire de recherche de l'ECE Paris-Lyon a évolué en termes de structures et d'axes de recherche.

En 2004, le premier laboratoire de l'ECE voit le jour, le LACSC (Laboratoire d'Analyse et Contrôle des Systèmes Complexes). A sa création, l'école commence à recruter des enseignants-chercheurs en fonction des besoins en enseignements. Les thématiques sont nombreuses et chacune est étudiée par un chercheur.

En 2014, un nouveau projet transversal est lancé qui est axé sur le véhicule du futur et plusieurs thématiques en informatique sont fusionnées. Puis en 2017 et dans le cadre du retour de la CTI (Commission du Titre d'Ingénieur), le centre de recherche est structuré en 3 axes à savoir nanosciences et nanotechnologies pour la santé et l'énergie, méthodes mathématiques pour l'ingénierie scientifique et systèmes intelligents communicants. La CTI considère la structuration comme un point positif, mais a recommandé une coopération entre les axes.

Enfin en 2018, le laboratoire lance un programme interdisciplinaire (PI-

ECE) axé sur la *smart city* qui se traduit par une extension du projet « véhicule du futur »

B) Partenariats

Le laboratoire a conclu plusieurs partenariats avec d'autres institutions de recherche ainsi que des établissements d'enseignement supérieur de renom. Parmi ces partenaires, nous pouvons citer le Laboratoire d'informatique Gaspard Monge, l'Institut des nanosciences de Paris (INSP), Télécom SudParis, l'Institut de recherche biomédicale (IRBA), le Politecnico di Milano, le Laboratoire d'ingénierie des systèmes de Versailles (LISV) ou

encore Institut VEDECOM. Tous ces partenariats sont en adéquation avec au moins l'un des axes de recherche du laboratoire de l'ECE.

III. Contexte du travail de recherche

Notre sujet s'intitule « **Prévision des défauts bancaires grâce aux états financiers** ». Afin de définir le sujet, il faut d'abord définir ses mots clés.

Un défaut bancaire est une insuffisance qui ne permet pas de satisfaire des obligations financières qui peut éventuellement mener à une crise bancaire. Dans le cadre de notre recherche, le défaut pourra être défini comme étant un déséquilibre financier dû à l'incapacité à satisfaire des obligations de la même nature.

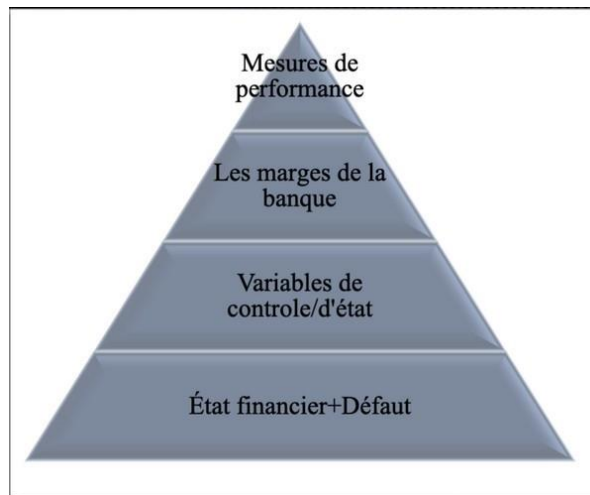
Un état financier est un ensemble de documents qui reflètent la santé financière d'une banque ou d'une entreprise. On y trouve des informations clés, notamment l'état de la structure financière de l'entreprise, la composition de son patrimoine, l'évaluation des performances et sa rentabilité.

En outre, l'analyse des états financiers d'une banque, pour prévoir un risque de défaillance, nécessite l'étude de plusieurs variables qui sont l'objet de notre étude, c'est-à-dire **les variables d'état** et **les variables de contrôle**. Les variables d'état sont celles qui nous déterminent le degré de réalisation de l'objectif tandis que les variables de contrôle sont celles sur lesquelles on peut agir pour faire évoluer le processus afin de mieux atteindre l'objectif. Elles seront abordées en approfondi dans une prochaine partie de ce rapport.

Par ailleurs, il existe une autre composante à intégrer dans notre analyse à savoir **le taux de marge**. Il correspond à la différence entre le taux d'intérêt auquel les banques prêtent et celui auquel elles se refinancent. Nous l'intégrons dans notre analyse car il constitue un indicateur montrant l'écart entre l'argent que la banque reçoit et celui que la banque émet.

Enfin, les analystes et les investisseurs peuvent, à partir de ces données, établir des **mesures de performance** qui déterminent les parts de marché de la banque, d'une part, et la survie de cette banque, d'autre part.

Afin de représenter la hiérarchisation de ces mots clés, nous vous présenterons cette pyramide qui montre l'importance de tous les termes que nous avons définis.



IV. État de l'art

Dans le cadre de notre recherche, nous avons fait un état de l'art afin de présenter les éléments déjà utilisés dans la détection des défaillances bancaires.

Variables de contrôle : d'après nos lectures, nous avons détecté 3 variables de contrôle à savoir le capital, la valeur des actifs et les liquidités. Le capital permet aux petites banques d'augmenter leur probabilité de survie et leur part de marché lors des crises, mais également aux moyennes et petites banques d'avoir une marge de négociation et de garanties sur le marché interbancaire ou bien lorsqu'elles auront besoin d'injections de liquidité de la part de la banque centrale.

Les banquiers affirment souvent que détenir plus de capital mettrait en péril leur performance et conduirait à moins de prêts, mais alors que les théories suggèrent une relation causale du capital à la performance,

nous reconnaissons qu'en pratique les deux pourraient être déterminés conjointement. En effet, le tableau 5 de Berger (voir annexe) montre comment les ratios de fonds propres bancaires d'avant la crise affectent la capacité des banques à survivre et à accroître leurs parts de marché au cours des cinq crises individuelles, tandis que le tableau 6 de Berger (voir annexe) examine trois canaux par lesquels le capital pourrait affecter la performance : la croissance des financements annexes, les prêts relationnels au bilan et les garanties hors bilan. Ces 3 éléments constituent ce que l'on appelle **variables d'état**.

La liquidité constitue notre deuxième et plus importante variable de contrôle. Elle correspond à la capacité d'une banque à obtenir des entrées d'argent liquide. Plusieurs facteurs expliquent pourquoi elle serait capitale à la survie des banques. Premièrement, les banques font partie d'un réseau qui les lie entre elles appelé marché interbancaire. Sur ce dernier, les banques créent un système de coassurance en cas de choc de liquidité. Pour cela, elles doivent disposer d'un système d'information symétrique, c'est-à-dire que toutes les banques devraient avoir le même niveau d'informations fournies. En pratique, ce n'est pas réalisable car chaque banque possède différents actifs. De plus, le système d'information asymétrique peut créer, lors d'une crise de liquidité, des inégalités de pouvoir interbancaire. En effet, les banques possédant plus de liquidités doivent en fournir aux banques qui en ont moins du fait du contrat de coassurance bancaire. Ce qui peut, potentiellement, assurer la survie des banques qui ne sont ni solvables ni efficaces, créant ainsi un état dans lequel ces banques ne ressentent pas le besoin de se réformer. Par ailleurs, les crises de liquidités ont montré que les banques possédant un excès de liquidité ont un monopole de pouvoir. En effet, elles ont un plus grand pouvoir de négociation face à des banques moins liquides et peuvent décider, pour des raisons de stratégie, de ne pas fournir de liquidité afin d'inciter les banques moins liquides à vendre leurs actifs au rabais, créant ainsi une répartition de ressources inefficace et parfois l'intervention de la banque centrale, soit pour offrir d'autres alternatives pour obtenir des liquidités, soit pour en injecter dans les banques en ayant besoin, ce qui peut inciter à une hausse des taux d'intérêts si le nombre de banques ayant besoin

d'injections de liquidité devient trop important. Ce qui pourrait restreindre les prêts bancaires.

La dernière variable de contrôle que nous avons trouvée est la valeur des actifs. En effet, ils jouent un rôle clé dans la survie des banques car certaines d'entre elles préfèrent conserver leurs actifs lors d'une crise bancaire puisque lorsque leur valeur augmentera dans le futur, cela assurera la stabilité de la banque. A l'inverse, les banques moins solvables auront tendance à vendre leurs actifs à un prix très réduit pour obtenir des liquidités le plus rapidement possible. Même si le prix des actifs se détermine généralement par l'offre et la demande, les banques peuvent également mal les évaluer, notamment quand ces actifs se trouvent en situation de bulle spéculative et qu'il y a un énorme écart entre la valeur intrinsèque de l'actif et celle de vente. Ce qui s'explique souvent par une politique monétaire et fiscale trop flexible qui incite les banques à choisir des actifs à fort rendement, accentuant ainsi l'effet de levier des banques qui, par conséquent, les rend plus exposées aux risques et donc plus fragiles. Pour cette raison, la valeur et la qualité des actifs est également importante à la survie d'une banque.

Mesures de performance : la plupart des investisseurs et analystes utilisent deux indicateurs pour mesurer la performance des banques à savoir les parts de marché et la survie d'une banque. La survie correspond à la capacité d'une banque à résister à un choc de liquidité lors duquel elle dispose de liquidités et de patrimoine pour assurer son existence pendant et après la crise. Nous voyons ici que la variable de contrôle qui impacte cette mesure est le niveau de liquidité. Quant aux parts de marché, elles correspondent au pourcentage de ventes totales dans une industrie générés par une entreprise particulière. Les parts de marché dépendent, en grande partie, de la taille de la banque et du capital dont elle dispose. Ce dernier peut également aux banques d'augmenter leurs parts de marché lorsqu'une crise financière éclate, mais que la banque dispose d'assez de capital pour avoir une plus grande marge de négociation.

V. Nouveautés de recherche

A l'issue de nos lectures, nous avons constaté que les chercheurs essayent d'expliquer l'origine des crises bancaires. Ces documents se basaient essentiellement sur la crise de 1930 et celle des *subprimes* qui s'est déroulée entre 2007 et 2009. Malgré la similitude des deux crises en termes d'impact et d'origine, nous constatons plusieurs caractéristiques propres à chacune constatées par les chercheurs.

L'étude de la crise bancaire de 1930 nous prouve que le secteur bancaire n'a pas uniquement subi la conjoncture internationale défavorable, mais présente aussi des faiblesses structurelles telles qu'une structure sectorielle fragmentée, des restrictions sur la disponibilité et le déploiement des capitaux, le manque d'infrastructure de soutien institutionnel, des lois du travail restrictives, une gouvernance d'entreprise faible, des pressions politiques et des réglementations inefficaces.

La notion de liquidité mesurée par les disponibilités et le portefeuille commercial est primordiale dans les schémas de faillites et liquidations durant la crise de 1930. Il existe une relation positive entre le niveau de liquidité et la probabilité de survie d'une banque.

L'analyse révèle une relation négative entre le niveau des acceptations et dettes interbancaires et la probabilité de survie. Bien que nous ne puissions pas conclure quant aux ressources propres et dépôts, le financement d'une banque par le marché interbancaire crée une vulnérabilité au risque de défaut.

Les banques avec un niveau d'influence plus élevé, mesuré par la distance moyenne pondérée variable mentionnée précédemment, sont plus susceptibles de recevoir de l'aide

Sans savoir si les banques diffèrent significativement dans leur allocation de portefeuille et leurs caractéristiques de niveau de prêt telles que le secteur, le type et la durée du contrat ou la partie destinataire, il est

difficile d'évaluer véritablement l'homogénéité des banques et d'identifier une contagion potentielle du portefeuille.

En analysant l'évolution du secteur bancaire français dans les années 1920-1930, on constate que la résistance des grandes banques provient des moindres retraits de dépôts auxquels elles ont dû faire face par rapport aux autres banques, de l'utilisation privilégiée de l'escompte au détriment des avances et d'une meilleure situation de liquidité au début de la crise qui leur a permis de mieux amortir les chocs.

Il existe un lien entre la probabilité de défaut d'une banque pendant la crise et les caractéristiques de son bilan. Le principal résultat qui se dégage de l'ensemble des modèles et des spécifications est l'impact négatif du ratio de liquidité sur la probabilité de défaut. Ce résultat est surtout valable lorsque nous restreignons la variable expliquée aux faillites et liquidations pour lesquelles l'exclusion des banques d'émission n'altère pas les valeurs et degrés de significativité.

En utilisant des t-tests à deux échantillons de Welch sur des données transversales de 1929, on constate également que les banques qui ont fait faillite ne diffèrent pas significativement de celles qui ont survécu en termes d'éléments clés du bilan et de ratios. Les banques qui ont fait faillite n'étaient pas nécessairement plus risquées, ni moins liquides, ni surallouées dans certains postes du bilan

La libéralisation financière correspond au processus de démantèlement de toute forme de contrôle réglementaire quantitatif ou qualitatif à caractère restrictif imposé par l'État sur les structures institutionnelles, les instruments et les activités des agents sur différents segments du secteur financier, non seulement au niveau interne, mais aussi à l'échelle internationale. Ceci est un facteur à l'origine des crises bancaires tout comme l'expansion du crédit. Ces principales causes évoquées conduisent à une hausse vertigineuse des valeurs mobilières (actions, obligations, entre autres) et immobilières, créant ainsi une bulle spéculative. Lorsque les prix de ces valeurs deviennent intenable, les investisseurs arrêtent l'achat de ces actifs, leurs prix chutent et le marché

s'effondre. Et comme les banques possèdent ces actifs qui se déprécient et manquent de liquidité, cela déclenche une crise de liquidité dans les banques. Ces dernières arrêtent de financer les entreprises qui, à leur tour, arrêtent la production, créant ainsi une récession économique.

Les crises bancaires succèdent des périodes d'inflation élevée ou de taux d'intérêt bas. 2 solutions peuvent remédier à ce problème : soit les banques centrales interviennent pour injecter de l'argent, ce qui permettrait aux banques privées de financer des projets à long terme, soit les banques centrales augmentent les taux d'intérêt, ce qui dissuaderait les banques de faire des emprunts illiquides.

L'expansion monétaire des banques vise à promouvoir la consommation qui, à son tour, engendrerait la croissance du PIB, mais ignore la potentielle inflation du prix des actifs qui pourrait survenir. Donc, il y a toujours un compromis entre la croissance économique et la stabilité du système financier.

Même si dans les 2 crises il y a un aspect de comportement irrationnel, la crise des *subprimes* comporte de nouvelles composantes qui la différencient de celle des années 1930 : prêts bancaires titrisés, politiques publiques visant à stimuler l'accès à la propriété, déséquilibres mondiaux (*global imbalances*), politique monétaire expansionniste visant à lancer la consommation des ménages et contrôle réglementaire faible.

VI. Nos contributions

Dans le cadre de ce projet de recherche, nous voudrions détailler l'ensemble de nos contributions qui vont constituer la feuille de route du second semestre.

Premièrement, nous créerons une bibliographie exhaustive pour ce travail de recherche en lisant non seulement les documents donnés par nos mentors, mais également d'autres lectures en rapport avec les crises bancaires. Il serait également enrichissant de lire des articles sur d'autres crises bancaires moins importantes comme celle de la Russie à la fin des années 1990. Cela pourrait nous apporter des informations concernant

les spécificités de chaque crise et les similitudes par rapport à la crise des années 1930 et celles des *subprimes*.

Puis, nous allons énumérer et étudier divers types d'information figurant dans les états financiers. A l'heure actuelle, nous ne savons pas ce qui se trouve dedans, mais nous tâcherons de choisir ces informations afin de les hiérarchiser et d'évaluer leur importance à l'heure d'un choc de liquidité.

Ensuite, nous établirons des relations entre les états financiers historiques, les faillites (pour les banques) et les rendements (pour les entreprises). Pour cela, nous utiliserons des modèles de *Machine Learning* que nous avons vues au début du premier semestre. Etant donné qu'il s'agit d'un problème de classification (apprentissage supervisé), nous utiliserons des techniques appliquées aux réseaux neuronaux comme le *forward propagation* ou le *backward propagation*. En termes de logiciel de programmation, nous utiliserons Python qui contient des bibliothèques adaptées au *Machine Learning* comme Numpy ou Scipy.

Enfin, nous interpréterons les résultats obtenus. Ceci nous permettra de créer un algorithme visant à prévoir les défaillances bancaires et de l'adapter en fonction des résultats obtenus.

VIII. Partie technique et développement

Notre travail a consisté, en premier lieu, en un ensemble de calculs statistiques permettant de vérifier s'il y a une significativité entre les banques qui ont fait faillite en fonction de leurs capitaux et leurs liquidités et les banques qui ont survécu. Cette étude a été faite avec les données de la période de l'entre-deux guerres (1918-1938).

Dans un premier temps, on a procédé à un t-test pour savoir s'il y avait une différence significative de moyenne entre les 2 groupes de banques (faillite et non-faillite). Néanmoins, les résultats obtenus ne nous ont pas permis d'avoir une conclusion claire sur le résultat. C'est pourquoi

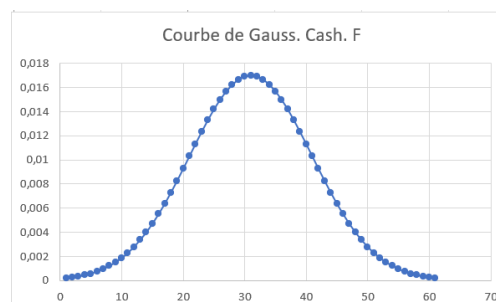
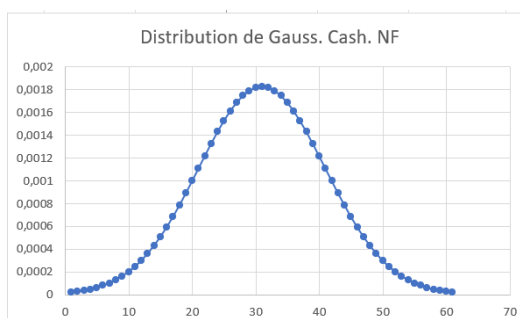
nous avons procédé au z-test. Concernant les banques qui ont fait faillite, on a conclu qu'en fonction du cash et du capital, la différence

entre les populations a été très significatives. Pour les banques qui ont survécu, on a aussi conclu qu'en fonction du cash et du capital, la différence entre les populations a été très significative. Ci-dessous les résultats des 2 tests statistiques qui ont été effectués :

| t-Test: Two-Sample Assuming Unequal Variances | | | t-Test: Two-Sample Assuming Unequal Variances | | |
|---|------------|------------|---|------------|-------------|
| | cap_nf | cap_f | | cash_nf | cash_f |
| Mean | 34.8384263 | 12.3369218 | Mean | 63.2672438 | 6.699273239 |
| Variance | 4575.42315 | 599.155786 | Variance | 47718.3279 | 398.8565421 |
| Observations | 138 | 219 | Observations | 138 | 219 |
| Hypothesized Mean Difference | 0 | | Hypothesized Mean Difference | 0 | |
| df | 160 | | df | 138 | |
| t Stat | 3.75593532 | | t Stat | 3.03407849 | |
| P(T<=t) one-tail | 0.00012072 | | P(T<=t) one-tail | 0.00144192 | |
| t Critical one-tail | 1.6544329 | | t Critical one-tail | 1.65597038 | |
| P(T<=t) two-tail | 0.00024143 | | P(T<=t) two-tail | 0.00288384 | |
| t Critical two-tail | 1.97490156 | | t Critical two-tail | 1.97730354 | |
| z-Test: Two Sample for Means | | | z-Test: Two Sample for Means | | |
| | cap_nf | cap_f | | cash_nf | cash_f |
| Mean | 34.8384263 | 12.3369218 | Mean | 63.2672438 | 6.699273239 |
| Known Variance | 4575.42315 | 599.155786 | Known Variance | 47718.3279 | 397.035279 |
| Observations | 138 | 219 | Observations | 138 | 219 |
| Hypothesized Mean Difference | 0 | | Hypothesized Mean Difference | 0 | |
| z | 3.75593532 | | z | 3.03411478 | |
| P(Z<=z) one-tail | 8.6348E-05 | | P(Z<=z) one-tail | 0.00120621 | |
| z Critical one-tail | 1.64485363 | | z Critical one-tail | 1.64485363 | |
| P(Z<=z) two-tail | 0.0001727 | | P(Z<=z) two-tail | 0.00241243 | |
| z Critical two-tail | 1.95996398 | | z Critical two-tail | 1.95996398 | |

A gauche, les résultats du t-test et du z-test du capital des banques qui ont fait faillite et qui ont survécu. A droite, les résultats du t-test et du z-test du cash pour les 2 groupes de banques

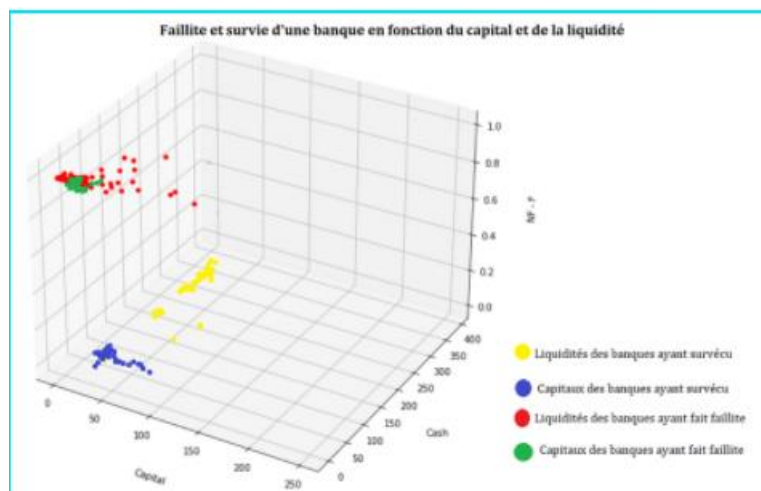
Nous avons, par la suite, décidé de faire 2 courbes de Gauss pour la liquidité des 2 groupes des banques.



La distribution des 2 courbes sont très similaires. Il y a donc une congruence entre les résultats des tests statistiques et les résultats du z-test et du t-test. Nous ne pouvons donc pas comparer les 2 courbes avec les résultats obtenus grâce aux z-test et t-test

Ensuite, nous avons classé les banques qui ont fait faillite et celle qui ont survécu en fonction de la moyenne des valeurs totales des actifs, car ce critère a été utilisé par Allen N. Berger (voir référence [2]).

Nous avons, par la suite, élaboré 2 modèles de *machine learning*, le réseau de neurones *feed-forward* et la régression logistique. En effet, il s'agit d'un problème de classification. Ceci veut dire que si la banque a fait faillite, on aurait 1 en sortie. Cependant, si la banque n'a pas fait faillite, le résultat serait 0. C'est ce que nous pouvons voir dans le graphe ci-dessous (en jaune, les banques qui ont survécu. En rouge, les banques qui ont fait faillite)



Représentation de la faillite et la non-faillite en fonction du capital et de la liquidité

Le résultat binaire des deux modèles sera obtenu grâce à la fonction sigmoïde. Si la variable X est supérieure à 0.5, alors la fonction sigmoïde nous donnera 1 en sortie. Sinon, la fonction nous donnera 0 en sortie

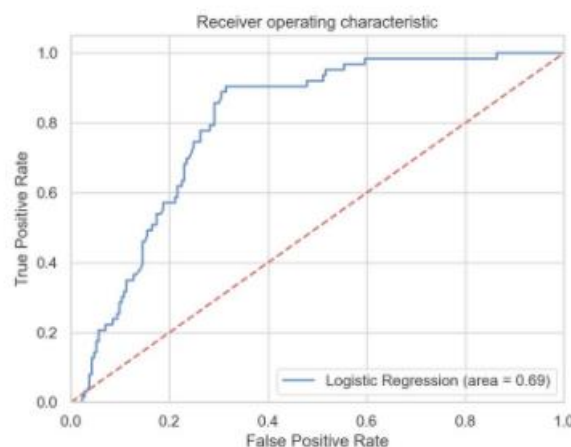
$$\text{Sigmoid}(\Theta X) = \frac{1}{1+e^{-(\Theta X)}}$$

Pour les 2 types de résolutions, les entrées seront le capital et la liquidité, toutes les 2 contenues dans une matrice de 2 colonnes. Nous ferons une matrice pour chaque groupe de banques (faillite et non faillite). Celle des banques qui ont fait faillite comportera, pour chaque banque, la troisième année avant la faillite et celle de la faillite. Pour le groupe des banques qui ont survécu, la matrice comportera le capital et la liquidité de 1934 et 1937. Les modèles seront ensuite entraînés avec 48% de l'ensemble des données. L'entraînement du modèle lui permettra d'apprendre par itération. Dans notre cas, nous en avons choisi 200. Une fois que les modèles ont été entraînés, nous les testerons avec des données dites test, c'est-à-dire 52% de notre ensemble de données, afin d'évaluer la précision de nos modèles et d'évaluer lequel des 2 est meilleur.

1) Régression logistique

La régression logistique est un modèle statistique utilisé principalement dans les problèmes de classification qui permet de mettre en rapport les valeurs d'entrées et les valeurs de sortie (x et y^{\wedge}). Comme toutes les analyses de régression, la régression logistique est prédictive.

A la suite de l'entraînement de notre réseau de neurones, nous avons obtenu une courbe qui nous permettra d'évaluer le taux de vrais positifs et de faux positifs.



Courbe ROC (receiver operating characteristic)

Nous observons que la courbe de régression logistique tend très à gauche. Cela veut donc dire que notre régression logistique est capable de détecter les banques qui ont fait faillite avec une précision de 69 %.

En outre, la matrice de confusion nous indique que ce modèle prévoit beaucoup de vrais positifs et un nombre non élevé de faux négatifs. Ceci est mis en évidence par sa matrice de confusion ci-dessous :

| Confusion Matrix-Logistic Regression | | | |
|--------------------------------------|--------------|---------------------|---------------|
| | | Predicted condition | |
| | | Positive (PP) | Negative (PN) |
| Actual condition | Positive (P) | 171 | 51 |
| | Negative (N) | 24 | 42 |

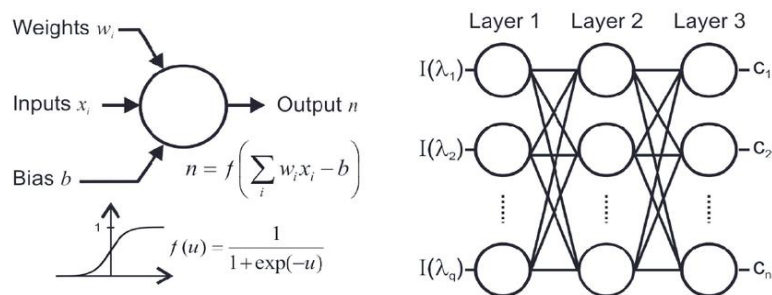
Matrice de confusion de la régression logistique

Nous avons, par la suite, élaboré un modèle de réseau de neurones pour évaluer ses performances et les comparer à celles de la régression logistique

2) Définition réseaux neuronaux :

Un réseau neuronal est une série d'algorithmes qui s'efforcent de reconnaître les relations sous-jacentes dans un ensemble de données par un processus qui imite le fonctionnement du cerveau humain. En ce sens, les réseaux neuronaux font référence à des systèmes de neurones, de nature organique ou artificielle.

Les réseaux neuronaux peuvent s'adapter à des entrées changeantes. Ainsi, le réseau génère le meilleur résultat possible sans avoir à redéfinir les critères de sortie.



Structure générale d'un réseau de neurones

Dans notre cas, notre réseau se compose d'un ensemble de couche. La première contient les neurones qui vont recevoir les 2 entrées de notre matrice. La deuxième est ce que l'on appelle une couche cachée qui comporte des neurones qui vont pondérer les résultats issus de la première couche. Enfin, c'est la troisième et dernière couche qui contiendra la sortie.

Nous avons utilisé les réseaux de neurones pour résoudre notre problème. Nous avons ajouté, dans la première couche, 80 neurones. Ensuite, nous avons créé une deuxième couche avec 40 neurones. La troisième et dernière couche contient le neurone qui nous donnera le résultat de la résolution du problème.

Puis nous l'avons testé sur des données dites test pour évaluer le modèle entraîné. Il s'avère assez précis. Cependant, il peut donner des faux négatifs, ce qui est très dans la prédiction des faillites comme l'illustre la matrice de confusion ci-dessous.

| Confusion Matrix-Neural Network | | | |
|---------------------------------|-------------|---------------------|--------------|
| | | Predicted condition | |
| | | Positive(PP) | Negative(PN) |
| Actual condition | Positive(P) | 194 | 23 |
| | Negative(N) | 52 | 12 |

Matrice de confusion du réseau de neurones

Les résultats des 2 modèles nous mènent à conclure que le modèle de régression logistique s'avère plus performant que le réseau de neurones dans la prévision des faillites. En effet, il est beaucoup plus précis, plus simple et donne moins de faux négatifs.

CONCLUSION

Nous avons élaboré deux modèles de *machine Learning*, le réseau de neurones et la régression logistique. Nous les avons utilisés en raison de la nature du problème, car la prévision des faillites ou de la survie des banques est considéré comme un problème de classification binaire.

Dans un premier temps, nous avons remarqué que, d'après nos lectures bibliographiques, le capital et la liquidité sont des facteurs significatifs qui jouent un rôle dans la survie des banques. En effet, le capital donne un large pouvoir de négociation aux banques qui en possèdent beaucoup plus, tandis que les liquidités sont un élément pour acheter des actifs en cours de dépréciation ou bien racheter d'autres banques qui font faillite. Cependant, nous avons remarqué que les banques qui ont fait faillite avait moins de liquidité que les banques qui ont survécu. Nous en déduisons donc que cette dernière variable a un poids prépondérant dans la survie d'une banque au cours d'une crise.

En ce qui concerne la partie technique, nous avons utilisés 2 modèles de *Machine Learning* pour résoudre ce problème de classification binaire à savoir la régression logistique et le réseau de neurones. Nous avons utilisé le capital et les liquidités comme données d'entrée pour prévoir la faillite ou la survie des banques. Cependant, il s'avère que le modèle de régression logistique est plus simple et plus performant. En effet, il est plus performant et donne moins de faux positifs que le réseau de neurones. Ce dernier modèle pourrait être amélioré en implémentant plus de couches neuronales et en augmentant le nombre de neurones ou le nombre d'itérations pour entraîner le modèle.

PERSPECTIVES

En termes de perspectives, nous souhaitons incorporer plus de données comme la valeur totale des actifs, la valeur totale des dettes ainsi que d'autres variables macroéconomiques de l'année de survie ou de faillite de chaque banque comme le taux d'intérêt ou l'indice des prix à la consommation. Nous souhaiterions également fabriquer d'autres modèles, comme des modèles bayésiens ou de *Deep Learning*, afin d'évaluer leurs performances dans la résolution d'un problème de classification.

REFERENCES

- Franklin Allen (2009) and Christa H.S. Bouwman. Financial Crises: Theory and Evidence (University of Pennsylvania). [1]
- Allen N. Berger (2013). How does capital affect bank performance during financial crises? (Journal of Financial Economics). [2]
- Nicole Fleskes (2020). Risk, Return, and Rescue: French Banks, 1920-1938 (Paris School of Economics). [3]
- Luc Laeven (2011). Banking Crises: A Review (Annual Review of Financial Economics | Home). [4]
- Edouard Vilpoux (2017). « La Crise Bancaire de 1930 en France : Une Nouvelle Analyse Historique et Quantitative des Données du Crédit Lyonnais » (Paris School of Economics). [5]
- PATRICE BAUBEAU, ERIC MONNET, ANGELO RIVA +, STEFANO UNGARO (2020). Flight-to-safety and the credit crunch: a new history of the banking crises in France during the Great Depression (The Economic History Review). [6]

ANNEXES

Table 5

Individual crises.

This table shows how precrisis bank capital ratios affect banks' ability to survive and enhance their market shares during the five individual crises (CRIS1: the 1987 stock market crash; CRIS2: the credit crunch of the early 1990s; CRIS3: the Russian debt crisis and Long-Term Capital Management bailout in 1998; CRIS4: bursting of the dot.com bubble and September 11; and CRIS5: the recent subprime lending crisis) and normal times (NORMALTIME) (see Section 3.1).

Results are shown for small banks [gross total assets (GTA) up to \$1 billion], medium banks (GTA exceeding \$1 billion and up to \$3 billion), and large banks (GTA exceeding \$3 billion), unless otherwise noted. GTA equals total assets plus the allowance for loan and the lease losses and the allocated transfer risk reserve (a reserve for certain foreign loans).

SURVIVAL is a dummy that equals one if the bank is in the sample one quarter before such a crisis started and is still in the sample one quarter after the crisis, and zero otherwise. Δ MKTSHARE, the percentage change in the bank's GTA market share is measured as the bank's average market share during a crisis minus its average market share over the eight quarters before the crisis, divided by its precrisis market share, and multiplied by one hundred. EQRAT is the equity capital ratio, calculated as equity capital as a proportion of GTA. The survival regressions are run only for cases in which there are a sufficient number of non-survivors. As a result, regressions are not run for medium and large banks during CRIS1, and for large banks during CRIS3. These cases are indicated as n/a.

To preserve space, we present only the coefficients on EQRAT. All regressions include the following common control variables: CREDIT_RISK, LOAN_CONCENTRATION, COMMERCIAL_REAL_ESTATE, CASH_HOLDINGS, SIZE, CORE_DEPOSITS, HQ_DEPOSITS, BRANCHES / ASSETS, and LOCAL_MKT_POWER. In addition, the small-bank regressions include BROKERED_DEPOSITS, TRADING_ASSETS, and MULTI_MKT_CONTACT; the medium-bank regressions include METRO_MKTS and MULTI_MKT_CONTACT; and the large-bank regressions include BROKERED_DEPOSITS and TRADING_ASSETS. Table 1 contains definitions of these variables. All dollar values are expressed in real 2010:Q4 dollars using the implicit gross domestic product price deflator.

t-statistics based on robust standard errors are in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

| Key variables | SURVIVAL | | | | | | Δ MKTSHARE | | | | | |
|-------------------------|--------------------|---------------------|------------------|---------------------|--------------------|--------------------|---------------------|---------------------|---------------------|---------------------|---------------------|--------------------|
| | CRIS1 | CRIS2 | CRIS3 | CRIS4 | CRIS5 | NORMAL TIME | CRIS1 | CRIS2 | CRIS3 | CRIS4 | CRIS5 | NORMAL TIME |
| <i>Small banks</i> | | | | | | | | | | | | |
| EQRAT | 20.841** (2.55) | 10.221*** (5.86) | 1.818 (0.62) | 10.405*** (4.79) | 5.601*** (3.95) | 9.556*** (4.27) | 0.962*** (22.17) | 1.138*** (19.41) | 0.215*** (4.94) | 0.739*** (13.09) | 1.176*** (30.64) | 0.266*** (7.01) |
| Number of observations | 10,755 | 9,811 | 7,798 | 7,146 | 6,340 | 15,447 | 10,755 | 9,811 | 7,798 | 7,146 | 6,340 | 15,447 |
| Adjusted R ² | | | | | | | 0.13 | 0.12 | 0.15 | 0.21 | 0.30 | 0.13 |
| <i>Medium banks</i> | | | | | | | | | | | | |
| EQRAT | n/a | 24.065** (1.91) | 29.094 (0.65) | 5.516 (0.69) | 0.512 (0.11) | 7.486 (0.82) | -0.102 (-0.54) | 0.776* (1.68) | -0.628 (-1.57) | -0.461 (-1.18) | 0.558*** (2.70) | 0.020 (0.11) |
| Number of observations | 275 | 283 | 270 | 238 | 323 | 574 | 275 | 283 | 270 | 238 | 323 | 574 |
| Adjusted R ² | | | | | | | 0.15 | 0.18 | 0.17 | 0.16 | 0.12 | 0.16 |
| <i>Large banks</i> | | | | | | | | | | | | |
| EQRAT | n/a | 65.647*** (2.19) | n/a | 14.956 (0.78) | 12.073 (1.13) | 26.419 (0.97) | 0.258 (0.33) | 0.988*** (2.22) | -1.518** (-1.98) | 0.780 (1.23) | 0.691* (1.94) | 0.397 (1.33) |
| Number of observations | 211 | 231 | 194 | 176 | 189 | 421 | 211 | 231 | 194 | 176 | 189 | 421 |
| Adjusted R ² | | | | | | | 0.14 | 0.09 | 0.05 | 0.01 | 0.12 | 0.04 |

Source : Franklin Allen (2009) and Christa H.S. Bouwman. Financial Crises : Theory and Evidence (University of Pennsylvania).

Table 6

Three channels through which capital may affect performance: noncore funding, relationship loans, and off-balance-sheet guarantees.

This table examines three channels through which capital could affect performance: growth in noncore funding, on-balance-sheet relationship loans, and off-balance-sheet guarantees. Panel A shows summary statistics on the three channels. Panel B regresses the three channels on precrisis capital ratios.

The crises include banking crises (BNCRIS: the credit crunch of the early 1990s and the recent subprime lending crisis), market crises (MKTCRIS: the 1987 stock market crash; the Russian debt crisis and Long-Term Capital Management bailout in 1998; and the bursting of the dot.com bubble and September 11), and normal times (NORMALTIME) (see Section 3.1).

Results are shown for small banks [gross total assets (GTA) up to \$1 billion], medium banks (GTA exceeding \$1 billion and up to \$3 billion), and large banks (GTA exceeding \$3 billion), unless otherwise noted. GTA equals total assets plus the allowance for loan and the lease losses and the allocated transfer risk reserve (a reserve for certain foreign loans).

NONCOREFUNDS is noncore funds (liabilities minus core deposits, the latter is the sum of transaction deposits, savings deposits, and small time deposits). RELSHIPLOANS is on-balance-sheet relationship loans (all loans excluding loans to depository institutions, foreign governments, and states). ORSGUARANTEES is off-balance-sheet guarantees (sum of loan commitments and commercial and standby letters of credit). Percentage changes (Δ) are measured as the bank's average value during a crisis minus its average value over the eight quarters before the crisis, divided by its precrisis value.

SURVIVAL is a dummy that equals one if the bank is in the sample one quarter before such a crisis started and is still in the sample one quarter after the crisis, and zero otherwise. Δ MKTSHARE, the percentage change in the bank's GTA market share is measured as the bank's average market share during a crisis minus its average market share over the eight quarters before the crisis, divided by its precrisis market share, and multiplied by one hundred.

EQRAT is the equity capital ratio, calculated as equity capital as a proportion of GTA. To preserve space, Panel B presents only the coefficients on the interaction terms although all the control variables (see Table 1) are included in the regressions. All dollar values are expressed in real 2010:Q4 dollars using the implicit gross domestic product price deflator.

t-statistics based on robust standard errors are in parentheses in Panel B. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

| Panel A: Summary statistics on the three channels | | | | | | | | | |
|---|---------------------|--------------------|--------------------|---------------------|---------------------|--------------------|---------------------|------------------|--------------------|
| | Small banks | | | Medium banks | | | Large banks | | |
| <i>Banking crises</i> | | | | | | | | | |
| %ANONCOREFUNDS (mean) | | | 0.183 | | | 0.128 | | | 0.082 |
| %RELSHIPLOANS (mean) | | | 0.144 | | | 0.149 | | | 0.145 |
| %ORSGUARANTEES (mean) | | | 0.854 | | | 0.484 | | | 0.373 |
| <i>Market crises</i> | | | | | | | | | |
| %ANONCOREFUNDS (mean) | | | 0.214 | | | 0.329 | | | 0.286 |
| %RELSHIPLOANS (mean) | | | 0.134 | | | 0.220 | | | 0.211 |
| %ORSGUARANTEES (mean) | | | 0.341 | | | 0.378 | | | 0.297 |
| <i>Normal times</i> | | | | | | | | | |
| %ANONCOREFUNDS (mean) | | | 0.268 | | | 0.312 | | | 0.284 |
| %RELSHIPLOANS (mean) | | | 0.146 | | | 0.200 | | | 0.222 |
| %ORSGUARANTEES (mean) | | | 0.355 | | | 0.340 | | | 0.320 |
| Panel B: Regression results | | | | | | | | | |
| Key variables | %ANONCOREFUNDS | | | %RELSHIPLOANS | | | %ORSGUARANTEES | | |
| | Small banks | Medium banks | Large banks | Small banks | Medium banks | Large banks | Small banks | Medium banks | Large banks |
| EQRAT*BNCRIS | 3.113*** (37.40) | 0.747*** (3.18) | 1.814*** (4.06) | 1.523*** (27.72) | 0.060*** (4.63) | 1.655*** (4.68) | 3.401*** (15.20) | 3.463* (1.93) | 3.846*** (3.49) |
| EQRAT*MKTCRIS | 1.975*** (24.82) | 1.191*** (2.80) | 1.825** (2.47) | 0.711*** (12.63) | -0.355** (-2.13) | 0.978** (2.41) | 2.257*** (10.62) | 0.773 (0.77) | 1.302 (1.06) |
| EQRAT*NORMALTIME | 1.404*** (11.60) | 0.702 (1.12) | 2.154*** (2.45) | 0.378*** (5.15) | 0.113 (0.36) | 0.705 (1.18) | 1.400*** (6.15) | 0.361 (0.29) | 1.292 (1.32) |
| Number of observations | 57,243 | 1,946 | 1,400 | 57,243 | 1,946 | 1,400 | 54,969 | 1,937 | 1,389 |
| Adjusted R ² | 0.20 | 0.24 | 0.18 | 0.20 | 0.19 | 0.10 | 0.17 | 0.14 | 0.11 |

Source : Franklin Allen (2009) and Christa H.S. Bouwman. Financial Crises : Theory and Evidence (University of Pennsylvania).