Une image contenant Police, Graphique, graphisme, logo

Description générée automatiquementUne image contenant Police, logo, Graphique, texte

Description générée automatiquement

POLYTECH TOURS

64 avenue Jean Portalis

37200 TOURS, FRANCE

Tél +33 (0)2 47 36 14 14

[www.polytech.univ-tours.fr](http://www.polytech.univ-tours.fr/)

Rapport de stage de 4e année (2023-2024)

**Exploration des grands modèles de langages pour réaliser des prédictions et détections d'anomalies dans des séries temporelles**

**Entreprise :**

EVERDYN

**Une image contenant Police, Graphique, capture d’écran, logo

Description générée automatiquement**10 Rue Aristide Briand  
37390 Notre Dame D’Oé  
France

**Tuteur Entreprise :**

FREMONT Thomas

Responsable Agence Centre Ouest

**Étudiant :**

MERCIER Titouan

Promo 2025

**Tuteur académique :**

RAGOT Nicola

# Sommaire

[Sommaire 1](#_Toc174175314)

[Table des illustrations 1](#_Toc174175315)

[Introduction 2](#_Toc174175316)

[1 Environnement 3](#_Toc174175317)

[1.1 Structurelle 3](#_Toc174175318)

[1.2 Relationnelle 4](#_Toc174175319)

[1.3 Contextuelle 5](#_Toc174175320)

[2 Synthèse du travaille effectuer 6](#_Toc174175321)

[2.1 État de l’art 6](#_Toc174175322)

[2.1.1 Familiarisation avec le Domain 6](#_Toc174175323)

[2.1.2 Fonctionnent des LLM 6](#_Toc174175324)

[2.1.3 Les techniques pour adapter un LLM 6](#_Toc174175325)

[2.2 Réalisation 6](#_Toc174175326)

[2.2.1 Faire un apprentissage : 6](#_Toc174175327)

[2.2.2 Évaluation des performances : 9](#_Toc174175328)

[3 Analyse réflexive de l’expérience 9](#_Toc174175329)

[3.1 Mon organisation 9](#_Toc174175330)

[3.2 Comparatif prévision/réalisation effectiv 11](#_Toc174175331)

[3.3 Les compétences acquises et consolidées 11](#_Toc174175332)

[Bilan de l’expérience 11](#_Toc174175333)

[Table des annexes 12](#_Toc174175334)

# Table des illustrations

**Aucune entrée de table d'illustration n'a été trouvée.**

# Introduction

Curieux et désireux d'évoluer dans ma pratique de l’informatique, je souhaite clôturer ma 4e année de formation d’ingénieur en réalisant un stage en entreprise. Ce stage me permettrait de pratiquer pour la première fois mes connaissance et compétence informatique au sein d'un environnement professionnel. Il sera également l’occasion d’avoir un potentielle aperçu de mon futur professionnel.

En effet je souhaite travailler dans le domaine de la cybersécurité. Pour tendre à cet objectif je me suis par ailleurs engagé l'année prochaine à poursuivre, en double diplôme, une maîtrise en cybersécurité. J'ai donc initié mes recherches de stage vers ce domaine afin d’acquérir ma première expérience en cybersécurité. Je suis également resté ouvert à toute autre proposition de stage dès l'instant quelle correspondait à ma formation d'ingénieur.

Cet état d'esprit m'a permis d'avoir l'opportunité de réaliser mon stage au sein de l'antenne tourangelle d’EVERDYN. Cette entreprise m’a permis de passer mes 9 semaines de stage autour du domaine de l'intelligence artificielle. En effet j'ai dû *Exploré les grands modèles de langages pour réaliser des prédictions et des détections d'anomalies dans des séries temporelles*.

Cela peut à première vue sembler être éloigner de mon désir professionnel. Mais à la vue de la proposition de celui-ci je me suis rendu compte que le domaine de l'IA met relativement obscur. En effet durant ma formation d'ingénieur mes choix de cours et de projets non pas eu l’occasion de tendre vers le domaine de l'IA. Au vu de l'engouement actuel et de son imbrication grandissent au sein de la société et des diffèrent domaine de l’informatique. Je trouvais très intéressant de passer 9 semaines à mieux le comprendre.

Annonce Plant

# Environnement

## Structurelle

Fondée en 2004, EVERDYN est une PME de 18 salarier domicilier à Aix en Provence mais présente dans tout la France grâce a sont réseaux de 5 agences. Ces agences ne sont pas réparties en centre de profit, il n’y a donc pas de concurrences entre elle.

EVERDYN évolue dans deux domaines d’activité. La première, leur principale activité présente au sein des 5 agences, est l’intégration, réparation et supervision de système. Ils sont capables de concevoir, programmer, démarrer la plupart des systèmes automatisés du monde industrielle.

Leurs seconds domaines d'activités, présent seulement au sein de l’agence d’Aix-en-Provence, est le développement de logiciels. Ils développent des logiciels industriels pour les utilisateurs d'AVEVA Software. Leur produit majeur est le logiciel et Interface de programmation d’application ALPANA. Elle permet de faire du « Dashboard ». Elle est capable de s'interface avec toutes les sources de donnée possibles et permet aux industrielles de monitorer leurs chaines de production et de partage est visualisé ces tableaux de bord sur tous types de support électronique.

Crée en 2007, L’agence de Tours où j'ai effectué mon stage travaille uniquement dans l'intégration, la supervision et réparation de systèmes. L'ensemble des collaborateurs de Tours sont des automaticiens.

Il travaille pour des vendeurs de matériel industriel. Leur rôle est, une fois le matériel livré à l'entreprise acheteuse, intégrer l’automatisation du nouveau système et gérer l’interfaçage avec le reste de la chaine de production. Dans ce cadre EVERDYN travaille régulièrement pour CIAM PAKE, spécialiste dans la conception de systèmes de convoyage de contenants. Actuellement ils travaillent ensemble à l'intégration d’un système de convoyeurs à l’entreprise Laiterie de Saint-Denis-de-l'Hôtel.

Ils ont également des clients pour des directs, c'est-à-dire qu'ils travaillent chez eux sur la supervision et le développement de l'automatisation de leur système. On peut citer par exemple l’entreprise SKF ou STMicroelectronics.

## Relationnelle

Durand mon stage j'ai majoritairement interagit avec 3 personnes.

M. FREMONT Thomas, mon maitre de stage :

Automaticien de métier, il est actuellement Responsable de l’agence de Tours. Cela signifie qu’il s'occupe de la gestion du personnel, de la planification des projets et de la répartition des effectifs chez les différends client. Cela occupe environ 10% de son activité. Le reste du temps, il exerce son métier d’automaticien et conseille éventuellement les autres collaborateurs grâce à l’expérience qu’il a pu accumuler depuis son arrivée au sein de la société en 2008.

Il a été là pour m’introduire dans l’entreprise, me fournir le matériel nécessaire à mon travail ainsi que répondre à mes différentes questions sur l’organisation de l’entreprise. Il a été également là pour vérifier que je n’étais pas en difficulté dans la réalisation de mon stage.

M. LE NY Nicolas, Gérant de EVERDYN :

Il travaille au siège social à Aix-en-Provence. C'est lui qui a défini mon sujet de stage et qui a contrôlé mon avancement. C'est également vers lui que je devais me tourner pour toutes mes questions relatives à la définition du besoin ou aux attentes de l’entreprise sur la réalisation de mon stage. Nous avons exclusivement échangé par email et téléphone.

Mme. LE NY Cécile, Assistente Administrative :

elle s’est occupée de mon accueil admiratif, vérifiant la validité de ma convention de stage et en s'assurant que j'ai eu à disposition tous les documents légaux obligatoires. Elle m’a aussi aider à obtenir toutes les informations à propos des politiques de qualité de vie au travail mi en ouvre au sein de la société.

Les différents collaborateur de l’agence de Tours :

Rarement dans les locaux de l’entreprise car souvent en déplacement chez les clients liée à la nature de leur travail. Ils sont nombre de 6, dons mon maitre de stage. J’ai eu le temps de les croiser tout au long de ma période de sage. Nous avons pu échanger lors de moments plus ou moins formels à propos de mon sujet de stage et de leur métier. (Il mon permis ...)

## Contextuelle

M. LE NY Nicolas a pour ambition de créer un outil informatique permettant d’anticiper les comportements des différents systèmes de l’industrie. Il permettra aux industriels de prédire les futurs dysfonctionnements des systèmes ou de prédire des variables lies à celui-ci comme par exemple la consommation énergétique ou le rendement.

Cet outil sera proposé aux entreprises souhaitant prédire le comportement de leurs systèmes sans avoir à investir dans la création d'un modèle de connaissance de celui-ci, qui peut s'avérer couteux.

Cet outil se veut générique et facilement adaptable afin d’être déployé dans tout type d'entreprise et pour tout type de système.

Pour fonctionner, cet outil se basera sur les données et covariables des systèmes sur lesquels il devra effectuer des prédictions et une détection d'anomalie sur les données produit par des systèmes. Dans l'industrie ces données sont nombreuses et déjà stockées de manière chronologique dans des bases de données. Cette solution devra donc être capable d'assimiler les différentes données des systèmes d’une entreprise spécifique afin que chaque solution vendue soit peaufinée pour l'acheteur. Cette solution devra également être capable de s'exécuter localement dans chaque entreprise, car la majorité des entreprises ne veulent pas que leur donner, soit exfiltrée de leur système. Les résultats de ces prédictions et détections seront ensuite utilisés par l'entreprise pour faciliter ces futures prises de décision.

Nous pouvons résumer cette solution en un mini cahier des charges de 2 fonctionnalités et une contrainte.

Mini cahier des charges :

FC1 : Prédire les futures valeurs d'une série temporelle

FC2 2 : Détecter une anomalie sur des valeurs de séries temporelles

CT1 : l'outil devra fonctionner en local au sein de l'entreprise pour lequel il fonctionne

Périmètre de mon stage :

L'objectif qui m’a été fixé par M. LE NY est d'explorer le potentielle de la technologie LLM pour la réalisation de prédiction et détection sur des séries temporelles afin de voir s'il y a un intérêt potentiel de l’utiliser pour le développement de cet outil. Son désir d'aller regarder vers cette technologie a été mû principalement par la lecture de l'article « Large Language Model Performance in Time Series Analysis » publié le 1er mai 2024 sur medium.com, un site de blog qui à une dimension de partade de contenue scientifique de par la présence de nombre d’articles sur divers type de technologie.

Il souhaitait également mieux comprendre cette technologie qui conné actuellement un grand engouement. ( faudra donc privilégier des LM peu volumineux capable de fonctionner sur de localement dans les entreprises. Ce qui a principalement pousser Monsieur le a envisagez d'explorer les solutions la M et la lecture de cet article … . Une autre contrainte est d'essayer de se tourner principalement vers des technologies françaises notamment celle de mistrale IA. Cela pourra faciliter l'acceptation de cette nouvelle technologie par les clients. )

# Synthèse du travaille effectuer

## Recherche sur le Domain

Durant les 2 premières semaines de mon stage j'ai commencé par étudier le sujet en réalisant un mini état de l'art. J'ai commencé par me renseigner sur les différentes techniques qui permettent de réaliser des prédictions sur des séries temporelles. J’ai résumer ces recherches dans la sous-partie suivante «  familiarisation avec le domaine ». Ensuite je me suis renseigné plus précisément sur la technologie LMM qui est au cœur du sujet de mon stage et également regardez les différentes méthodes existantes pour adapter leur comportement à une tâche spécifique. J’ai retranscrite c’est résulta dans les partie 2.1.2 et 2.1.3

### Familiarisation avec le Domain

La prédiction sur des valeur à partir de série temporelle n’est pas un sujet nouveau si les rossant avance en intelligence artificielle ont permis des progrès notamment en pouvant prendre en conte un nombre élargie des facteurs qui pouvaient peuvent influencer le comportement d'une série temporaire. Les premiers travaux de recherche à ce sujet ont été initiés dans le domaine de la statistique et de l'analyse de données.

### Fonctionnent des LLM

Les LLM, grands modèles de langage, sont des systèmes d'intelligence artificielle. Ils sont majoritairement utilisés dans le cadre du traitement du langage naturel. En effet les LLM sont conçu pour le traitement de donnée séquentielle. Ils sont entrainés sur de très grand Corpus de texte afin d’être en capaciter de prédire la probabilité d’une séquence de mots en fonction des mots précédant. Ce fonctionnement permet de les utilisées pour de nombreuse application, par exemple la traduction de texte, Résumé de texte, etc. C’est notamment cette technologie qui a permis l’arrivée d’outillé comme tchat gpt.

C’est LLM fonctionne grasse a une architecture de réseaux de neurones profond, et plus particulièrement sur le récente modèle d'architecture « Transformer ». En effet ce type de modèle a était introduite en 2017 dans l’article « **Attention Is All You Need** » comme son nom l'indique les Transformers fonctionnent uniquement avec un mécanisme d'attention. Contrairement aux précédente technologie de réseau de neurone …

Le mécanisme d'attention permet de sélection les parties importantes en attribuant des poids aux différents tokens un l'entrée du réseau de neurones

ce qui donne comme avantage …

Pour résumer plus simplement voici les différentes étapes qu'il se passe lorsque l'on fait une requête dans un LMM :

1. Le tesxe et découper en petite partie appeler token
2. C’est toquenne son convertien en vecteur de

### Les techniques pour adapter un LLM

## Réalisation

### Faire un apprentissage :

Après des recherche moment de recherche et plusieurs tentatives, j’ai pris la décision d’utiliser l’environnement Hugging Face.

Exécuter un modèle Hugging Face avec Ollama

### Création des datasheets

### Évaluation des performances :

Hé hé afin d'évaluer les performances des prédictions j'ai comme nous l'avons vu dans l'état précédente à chaque apprentissage gardez une partie des datasets comme base de validation et donc pour ne pas avoir hé le problème de… cette base de validation n'était pas utilis dans le processus d'apprentissage. Hé donc j'ai donc créé un script python qui compare Hey de série temporel lune la série temporelle réelle et l'autre générée par h le LM.

Pour réaliser cette comparaison je crée un graphique j'ai affiché les 2 courbes afin de pouvoir de manière visuelle rapidement R si les prévisions sont plutôt bonnes ou mauvaises. Hé je calculais également différentes métriques. Je récupérer la distance quadratique moyenne. La plus grande distance quadratique est la plus petite. Je calculais ces métriques 2 fois une fois sur les séries temporelles normales et la 2nde sur les séries temporelles mais en les ayant normalisées.

Je veux calculer également une métrique appelée hé performance de la prédiction. Hé cette valeur est un pourcentage hé 100% correspond à une prédiction exacte et le seuil de 0 pour-cent était atteint lorsque la prédiction 2 fois supérieur à la valeur réelle.

Résulta :

Évaluation des performances des différente prédiction est détection réaliser pendant le stage :

Premier aisser : Dans un premier temp j’ai effectué des test avec les version gratuit disponible en ligne des LLM Claude, Mistral, GPT3

Résultat du premier apprentissage

Résultat premier apprentissage avec le modèle mistral7B :

# Analyse réflexive de l’expérience

## Mon organisation

Durant mon stage j'ai dû m’organisation seul pour planifier les différentes tâches à effectuer pour aboutir résolution de mon sujet. J’ai mis en palace différente technique, que nous verrons si après, pour être efficace et contrôler mon avancement. Dans cette mise en palace j’ai dû prendre en compte seule de contrainte organisationnelle imposée par l'entreprise. À savoir : Les Horaire de travail, êtres présent 7 h par jour dans les locaux de l’entreprise du lundi au vendredi et la rédaction d'un Rapports d’activité hebdomadaire.

Les outils et méthode de mon organisation :

**Tableau de Planification :** Dès la première semaine, j’ai découpé l’entièreté de ma période de stage en période ponctuer par des datte clé. Vous pouvez retrouves cette planification en annexe B. le rôle de cette planification est d’avoir une route à suivre et d’avoir une vision sur le moyen terme du dérouler de mon stage. Cette planification n’est pas rigide elle peut évoluer aux files des imprévue et apparition de nouvel évènement. Nous dans d’ailleurs dans la sous partie suivent « Comparatif prévision/réalisation effective »

**Liste des taches à effectuer :** En complément du Tableau de Planification j’ai créé une liste de tache sur le bloc note de mon téléphone. Je mettais à jour cette liste de façon hebdomadaire. Cela ma permis d’avoir une vision sur le court terme des différente tache à effectuer.

**Le Week :** C’est un rapport hebdomade sous la forme d’un PowerPoint, vous pouvez le trouver en Anexe A. chaque employer, stagiaire et apprenti de l’entreprise doit le remplir et l’envoyer a sont supérieure. La fonction de se rapport et de suivre l’activité des salarie et faire remontes d’éventuelles difficultés qui pourrais apprêtes. Par exemple grâce à la section « Problème » ou « retours QHSE». Il permet également à l’employer de structurer l’excise de son activité, notamment à travers les deux section « réalisation » et « prévision ».

**Cahier de stage :** J’ai choisi affin d’avoir une persistance dans les réflexions et information importent que je peut rencontrer dans mon stage de réalise un genre de « Caille de bord » ou je note de fassent chronologique et séparent par jour l’ensemble de mes note que j’ai pu être amener a prendre pendants mon stage. Celas permet de facilite le retour rétrospectif et de retrouver rapidement une information dons on aurais besoins et qui aurai étais prise à un moment éloigner de son utilisation.

**Historique de recherche :** sout forme d’un document texte je sauvegarder tous les lien des site internet que j’ai trouvé pertinent pour mon travail. Je rédiger une petite description pour chaque lien afin de les distinguer fsilement entre eux. Le rode ce document de garder une trace de mon travail de rêcherche. Et ainsi pouvoir retrouve facilement des élément qui pour mètre de nouveaux utiles.

**Git  :** j’ai utilise cet outil de vertionigne fonctionent avec un dépôt distant pour sauvegardes l’ensemble de mes production écrite ( code, dataset, divers document explicatif, rapport de stage …). J’ai utiliser cette outile pour ne pas avoir de perte de donnée et faciliter ma mobilier de travail, car j’ai été amener a travaillé sur plusieurs pc diffèrent.

Pour ce stage l’utilisation de ces outil et météo a bien fonction, sont avantage est que l’ensemble des outils sont simple et peut contraigne, ce qui fait que le temps passe à l’organisation est très faible par rapport eau temps passée a la réalisation des différente tache.

Mon Historique de recherchem’a étais très utile pour la partie « 2.1 État de l’art » que j’ai rédigée à fin de mon stage environ 6 semaine après la fin de mes recherche. Il aurait étet plus facille de redijait au mois un eboche a la fin de ma fase de rechercher je pense que j’aurrai passe mon de temp a me rememorais setaine partie et.

Ici mon durant ma peride de stage mon travaille ne ma pas amene a collaborer donc avec d’autre personne

La planification :

Au cours de ma première semaine, j’ai constaté que j'étais laissé en grande autonomie pour réaliser mon stage. J’ai décidé de mettre en place une planning afin d'avoir différentes étapes à suivre pour aboutir à la réalisation du sujet. Cette planification c’est découper en 4 grande étape :

La premier étape « familiarisation avec le sujet », je l’ai fixé à une durée de 2 semaine. Cette première période devait me permettre de mieux comprendre les différentes technologies du domaine de l’IA et de comprendre plus en détail le fonctionnement des LLM.

La second étape « choix des techno », d’une durée de 1 semaine. Cette période a pour rôle de choisir les technologies et étape à suivre pour adapter un LLM.

La 3e étape «  réalisation des apprentissages », d'une durée de 4 semaines. Cette étape est l'étape centrale du stage elle permet de réaliser des apprentissages et pour parvenir à cela j'ai réalisé de manière itérative les

3 sous étapes suivantes : h réalisation des datasets pour un apprentissage hé c'est au cours de cette étape qu'on choisit comment structurer les données en fonction de la stratégie d'apprentissage choisie la semaine précédente, réalisation de l'apprentissage c'est là où mais en place le modèle on choisit les différents paramètres de l'apprentissage p et enfin la dernière test du modèle c'est là où on évalue les performances du modèle et où peux faire de nouveaux choix avant de réitérer ces 3 étapes.

La 4e étape de clôture, d'une durée de h 2 semaines. Que cette étape a pour rôle 2 rédiger mon rapport de stage mettre en forme hé les travaux que j'ai pu effectuer ainsi que réaliser une synthèse des résultats pour hé hé h ppouvoir les retransmettre à l'entreprise de manière écrite ou également lors d'une présentation

## Comparatif prévision/réalisation effective

## Les compétences acquises et consolidées

Développement python. Petp8. Librerie Transformeurs.

Environnement, conda

GIT LFS

# Bilan de l’expérience

C'est 9 semaines de stages ont été très enrichissantes et m'ont permis de m'ouvrir au monde de l'intelligence artificielle qui était jusque-là plutôt obscur. Hé cela m'a permis de comprendre les différents technique appliquer à ce domaine et de pratiquer l'une d'elle les LLM

hé, mais j'ai également pu développer des compétences moins technique blabla

hé hé cette expérience aura donc certainement une influence très positive sur mon futur professionnel… h

Et très positive j'ai permis de très développer mon autonomie dans la gestion de projet et dans son aboutissement, mais bon j'aurais aimé quand même avoir puis pouvoir échanger qualifié dans mon domaine de stage avec afin d'obtenir des conseils/guider techniques afin d'accélérer ma réalisation et d'augmenter on apprentissage.

Ces 2 semaines de stage n'ont pas changé mon projet professionnel, mais m'a permis comme je le souhaitais avant d'arriver de me familiariser avec un nouveau domaine et tester des technologies de celui-ci.

# Table des annexes

A - Organigramme page : 19

Source : Intranet de l’entreprise

**Exploration des grands modèles de langages pour réaliser des prédiction et détection d'anomalies dans des séries temporelles**

**Résumé :**

Durant 9 semaines j'ai évolué au sein d'une entreprise d'intégrateur industriel qui ont une seconde  activité dans le développement de logiciel.  J'ai participer au prémisse de  leur projet de développement d'un outils de prédiction et détection d'anomalies dans des série temporelle.

Ma mission a été de tester et d’évalué le potentielle de la technologie LLM pour réaliser cette outil. Cela ma permit de voir les différente techno de LLM disponible en open source. Comprendre et réaliser un apprentissage sur un LLM afin de l’adapter à un tache particulière, sur un jeu de donnée particulier. Puit évaluer et comparer les performance de ces LLM.

Cet outil sera à destination des industriels qui souhaitent avoir de prédictions sur le valeur de retours de leurs systèmes en fonction des covaroble qui peuvent l'influencer.

**Mots-clés :**

3 à 8 mots-clés

LLM :

Série temporelle

**Abstract :**

*5 to 15 lines to provide an overview of your internship work*

Keywords :

3 to 8 keywords