

Datathon 2025



POLYFINANCES



SOMMAIRE

1. Challenge.....	03
Contexte.....	03
Description.....	03
Démarche proposée (guide).....	04
Extraction des informations.....	04
Modélisation de l'impact réglementaire.....	04
Simulation et recommandations au niveau portefeuille.....	05
Conseils.....	05
2. Données.....	06
3. Livrables attendus.....	08
Soumissions DevPost.....	09
Soumissions Courriel.....	09
Recommandations.....	10
Démonstration vidéo.....	10
Documentation technique.....	12
One Pager.....	13
4. Critères d'évaluation.....	15



1. CHALLENGE

Contexte

Vous évoluez dans un monde où les marchés financiers sont influencés par un cadre réglementaire de plus en plus complexe, une intensification des lois protectionnistes et de différentes sanctions économiques. Ces règles redéfinissent la gestion d'actifs et forcent les investisseurs institutionnels à être plus agiles et plus résilients dans leurs stratégies.

Votre défi est d'imaginer comment transformer cette complexité réglementaire en opportunité d'aide à la décision pour la gestion de portefeuilles d'actions.

Description

Le défi consiste à concevoir un outil d'analyse et de simulation capable d'exploiter les différentes techniques d'IA générative pour analyser automatiquement de nouveaux textes réglementaires.

Votre solution devra en extraire les éléments clés, en évaluer l'impact potentiel sur un portefeuille d'actifs de référence, le S&P 500, puis transformer ces résultats en recommandations stratégiques.

Pour y parvenir, vous serez amenés à croiser ces informations avec des données d'entreprises et de marché afin de :

- simuler différents scénarios
- mesurer les risques pour les entreprises, secteurs ou zones géographiques et identifier les entreprises les plus à risque d'être impactées
- proposer des ajustements concrets (réallocation, rotation sectorielle, remplacement de titres, etc.).

L'objectif est de démontrer comment l'IA peut transformer la veille réglementaire en un outil proactif d'aide à la décision, directement exploitable dans la gestion de portefeuilles.

Il faut que l'outil ne soit pas limité à un seul type de document. Peu importe la structure ou le format du texte (loi, rapport, etc.), il devrait pouvoir s'adapter facilement. Montrer cette flexibilité sera un véritable atout.



Démarche proposée (guide)

Extraction des informations

Utilisez une combinaison de techniques de traitement automatique du langage (NLP) et de modèles d'IA générative pour extraire :

- les entités concernées (entreprises, secteurs, pays)
- le type de réglementation
- les dates d'application
- les mesures imposées, etc

Allez plus loin en croisant ces textes réglementaires avec les informations issues des documents de référence des sociétés, par exemple les rapports annuels déposés sous la forme de document 10-K pour comprendre leur modèle économique.

Exemples: Apple — fournisseurs Foxconn, Corning, TSMC, etc.

Modélisation de l'impact réglementaire

- À partir des données extraites, évaluez l'impact attendu sur les actions individuelles.
- Votre outil devra au minimum fournir un score de risque ou d'impact par entreprise.
- Vous pouvez aussi fournir des résultats chiffrés (perte estimée, % d'exposition).
- Mettre en évidence les explications sur votre raisonnement pour renforcer la transparence de votre outil.

Par exemple, une taxe de 100 % sur les semi-conducteurs aux États-Unis n'aurait pas d'impact direct sur les profits d'Apple, puisque ses téléphones sont fabriqués en Inde et en Chine. En revanche, les tarifs appliqués sur les téléphones augmenteront la pression sur les marges d'Apple pour les ventes d'iPhone aux États-Unis.



Démarche proposée (guide)

Simulation et recommandations au niveau portefeuille

- Agrégez les impacts pour évaluer l'effet global sur le portefeuille de référence (S&P 500).
- Identifiez les zones de concentration du risque réglementaire (par secteur, par pays, etc.).
- Proposez des ajustements qualitatifs ou quantitatifs (rotation sectorielle, remplacement de titres, réallocation géographique).
- Concevez une interface web interactive permettant de visualiser clairement l'exposition du portefeuille et les ajustements proposés.

Conseils

- Ces étapes servent de point de départ. Vous pouvez les modifier, les améliorer ou explorer d'autres approches.
- Faites preuve de créativité et d'originalité : la valeur de votre solution ne réside pas seulement dans la technique, mais aussi dans la pertinence des recommandations et la clarté de votre présentation.
- Soyez attentifs à l'expérience utilisateur : plus votre outil sera simple à utiliser et à comprendre, plus il marquera positivement le jury.



2. DONNÉES

Dans le jeu de données, vous trouverez plusieurs fichiers et dossiers contenant l'ensemble des informations nécessaires à la résolution du cas.


Il est également autorisé de consulter et d'utiliser toute information pertinente provenant de sources externes, telles que Yahoo Finance, Morningstar, ou encore le site officiel du gouvernement américain pour la publication des rapports de la SEC (Securities and Exchange Commission), notamment les rapports 10-K et 10-Q. Lien du site : <https://www.sec.gov/edgar/search/?r=el#>

IMPORTANT : Un nouveau document, qui sera pris en compte pour l'évaluation finale, vous sera transmis dimanche matin.

Commencez par tester votre application sur un échantillon restreint de données, puis conservez les résultats d'extraction afin de limiter les appels répétés aux services AWS et d'optimiser leur utilisation.

2025-08-15_composition_sp500.csv : liste de composition du S&P 500 au 15 août 2025. Ce fichier contient des colonnes comme le ticker, le nom de l'entreprise, le poids du ticker dans le S&P 500, et le prix courant de l'action.

2025-09-26_stocks-performance.csv : performances des actions au 26 septembre 2025. Ce fichier contient un enregistrement par ticker avec des métriques de performance (valorisation boursière, revenu net, EPS, FCF, etc.).



Dossier directives : Contient des documents législatifs et réglementaires (HTML ou XML) provenant de différentes juridictions (Canada, États-Unis, Japon, Europe, Chine). Ces fichiers sont des textes officiels (règlements, lois, projets de loi). Le fichier README.MD dans le dossier explique la provenance de chacun des documents législatifs.

- DIRECTIVE (UE) 20192161 DU PARLEMENT EUROPÉEN ET DU CONSEIL.html
- H.R.5376 - Inflation Reduction Act of 2022.xml
- REGULATION (EU) 20241689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL.html
- 中华人民共和国能源法_中国政府网.html` [loi chinoise sur l'énergie]
- 人工知能関連技術の研究開発及び活用の推進に関する法律.html` [loi japonaise relative à l'IA]

Dossier fillings : Le dossier contient un dossier par ticker (par exemple 'AAPL/', 'MSFT/', etc.). Chaque ticker contient le rapport annuel de 2024 de l'entreprise qui lui correspond.

Le rapport 10-K est un document officiel que toutes les entreprises cotées en bourse aux États-Unis doivent déposer chaque année auprès de la Securities and Exchange Commission (SEC). Ce rapport offre une vue détaillée des activités, risques, situation financière, dépendances régionales et chaînes d'approvisionnement des entreprises, ce qui permet d'identifier précisément leur exposition et leur vulnérabilité face à de nouvelles réglementations.

Voici quelques dépôts GitHub intéressants pour l'extraction et l'analyse automatique de rapports financiers (10-K, 10-Q, etc.) :

- alphanome-ai/sec-parser — Outil permettant de parser et de structurer automatiquement les rapports financiers de la SEC à partir des fichiers EDGAR — <https://github.com/alphanome-ai/sec-parser?tab=readme-ov-file>
- janlukasschroeder/sec-api-python — Client Python pour interagir facilement avec l'API de la SEC et extraire des données financières ou textuelles — <https://github.com/janlukasschroeder/sec-api-python>
- DrishtiiAggarwal/10-k-10-q-extraction — Exemple de pipeline d'extraction et d'analyse de contenu à partir des rapports 10-K et 10-Q — <https://github.com/DrishtiiAggarwal/10-k-10-q-extraction>



3. LIVRABLES ATTENDUS

⚠ **Respectez les heures de remise: Fuseau horaire - Montréal / ET (America/Toronto)**

Tout manquement aux règles entraîne une disqualification immédiate.

Pour chacun des livrables à produire, une liste d'exemples issus des éditions précédentes est mise à votre disposition dans le drive qui sera envoyé dans le canal ressources du Discord. Ces exemples, issus de l'édition précédente, sont fournis uniquement à titre de référence de format. Ils ne correspondent pas au sujet de cette édition 2025 et ne doivent en aucun cas être repris tels quels dans vos rendus. Veillez à produire un contenu original et adapté spécifiquement à votre solution au défi de cette année.

La soumission se déroulera en deux phases, afin de laisser plus de temps pour remettre la vidéo de démonstration ainsi que les autres livrables.

- **Pour la première phase**, les livrables devront être déposés sur la plateforme Devpost à l'adresse suivante : <https://datathon-polyfinances-2025.devpost.com>. Si vous n'avez pas encore de compte Devpost, vous devez d'abord en créer un, puis rejoindre la compétition Datathon 2025 Polyfinances. Une fois inscrit, formez ou rejoignez votre équipe directement sur le site.
- **Pour la seconde phase**, les livrables seront à envoyer par mail. Voir les informations ci-dessous.



Soumissions DevPost : Avant dimanche 2 novembre – 22:59

LIEN DEVPOST : <https://datathon-polyfinances-2025.devpost.com>

1. Outil déployé en ligne : Mettre l'URL de l'outil en ligne dans l'espace prévu à cet effet
2. Code source : Mettre l'URL du dépôt GIT de l'outil en ligne dans l'espace prévu à cet effet
3. Ajoutez un bref guide de démarrage de l'application dans le README du dépôt.

⚠ Rappel AWS : Accès jusqu'au lundi 3 novembre 13:59, mais aucun changement sur l'outil après dimanche 2 nov. 23:00. Toute modification après ce délai = disqualification immédiate.

Soumissions courriel : Avant lundi 3 novembre – 14:00

Envoyer par e-mail (canal officiel) :

À : logiciel@polyfinances.ca

Sujet : Datathon - Équipe XX

1. One pager Nom de fichier : TEAM-XX_onepager.pdf
2. Démo vidéo [durée maximale 5 minutes] [accès Google Drive] Nom de la vidéo : TEAM-XX_video_datathon_2025
3. Document technique Nom de fichier : TEAM-XX_doc_tech.pdf
 - Ce schéma doit donner au jury une vue d'ensemble claire de votre solution et de son intégration à l'environnement AWS. Contenu attendu :
 - Schéma d'architecture (composants, flux, services AWS, coûts estimés si pertinent)
 - Choix technologiques et justifications



Recommandations

Démo vidéo (durée maximale : 5 minutes)

⚠ Rappel AWS : Accès jusqu'au lundi 3 novembre 13:59, mais aucun changement sur l'outil après dimanche 2 nov. 23:00. Toute modification après ce délai = disqualification immédiate.

1. Introduction et présentation de la solution (≈ 1 min)

- Présentation générale de l'outil créé
- Démonstration des différentes fonctionnalités et de leur rôle.
- Explication de l'objectif global de l'outil.

2. Architecture technique AWS (≈ 1 à 1 min 30)

- Explication de l'architecture mise en place sur AWS, en s'appuyant sur le document technique.
- Description de la pipeline : étapes clés et interactions entre les différents composants.

3. Mise en pratique : analyse d'un document juridique (≈ 2 min)

- Mise en évidence des insights extraits du texte et de leur classification/utilisation.
- Illustration de la valeur ajoutée : impact direct sur les recommandations et sur le portefeuille.
- Présenter un maximum de résultats et de sorties, tout en expliquant clairement les décisions prises et la logique.
- Montrer que votre solution est adaptable peu importe les données fournies [loi / directive / rapport]

4. Critiques et pistes d'amélioration (≈ 30 sec)

- Retour sur les limites actuelles de l'outil.
- Conclusion rapide.
- Suggestions d'améliorations possibles [techniques ou fonctionnelles] [ouverture]



Recommandations

Démo vidéo - Conseils pratiques

1. Anticipez et soignez la réalisation.

- La préparation d'une vidéo demande plus de travail qu'il n'y paraît. Prévoyez du temps pour tourner et monter la vidéo. Une démo précipitée ou incomplète peut réduire considérablement l'impact de votre projet.

2. Préparez un script clair et structuré. Même si vous ne le suivez pas mot pour mot, définissez un fil conducteur avec un minutage précis. Identifiez à l'avance:

- Les fonctionnalités principales à démontrer.
- L'élément accrocheur à présenter (fonctionnalité phare ou cas d'usage marquant).
- Une conclusion forte qui met en avant l'impact concret ou le bénéfice utilisateur.

3. Montrez la solution en action. Le jury doit saisir le fonctionnement de votre solution à travers son utilisation réelle.

- Illustrez un cas d'utilisation concret du début à la fin.
- Pour éviter les imprévus, privilégiez l'enregistrement de séquences testées plutôt qu'une démonstration en direct.

4. Assurez une qualité visuelle et sonore.

- Utilisez un micro de bonne qualité.
- Ajoutez des annotations ou légendes brèves pour guider la compréhension.
- Optez pour des visuels sobres et clairs: zooms sur les moments clés, transitions fluides, pas de surcharge graphique.

5. Marquez les esprits. Votre objectif est de rendre la vidéo mémorable. Posez-vous la question : qu'est-ce qui retiendrait l'attention d'une entreprise ou d'un investisseur?

- Commencez avec un hook fort (statistique frappante, phrase d'accroche, mise en situation).
- Montrer des résultats concrets: gain de temps, simplification des tâches, insights générés.
- Terminez par une conclusion impactante.



Recommandations

Documentation technique

Présentez les différents composants (interface utilisateur, backend, moteurs d'analyse, bases de données, API externes, etc.) ainsi que les flux d'information entre eux. L'objectif est de montrer comment les éléments s'articulent.

Choix technologiques : incluez les frameworks, bibliothèques et outils que vous avez retenus et démontrez visuellement leur rôle dans votre solution.

Services AWS : identifiez clairement les services AWS que vous avez mobilisés.

Exemples: S3, OpenSearch, Comprehend, Bedrock, etc.



Recommandations

One pager

Le One Pager est un outil stratégique pour présenter votre solution de manière claire, concise et impactante, en mettant rapidement en évidence sa valeur ajoutée auprès du jury.

- **1 page maximum** : mise en page simple et visuelle pour faciliter la lecture.
- **Hierarchiser l'information** : titres clairs, phrases courtes, mots-clés en évidence.

Contenu

1. Titre de la solution

- Choisissez un nom professionnel, accrocheur et représentatif de votre solution, afin que le jury puisse facilement l'identifier et s'en souvenir.

2. Contexte et description de la solution

- Problème : Décrivez de manière concise le défi spécifique auquel sont confrontés les analystes financiers et que votre solution cherche à résoudre.
- Solution : En 1 à 2 phrases, expliquez comment votre solution répond à ce problème

3. Fonctionnalités clés

- Énumérez les fonctionnalités principales de votre solution.
- Pour chaque fonctionnalité, décrivez ce qu'elle apporte à l'utilisateur, en mettant en avant les avantages concrets et les éléments innovants.

4. Technologies et ressources utilisées

- Technologies AWS : indiquez les services spécifiques que vous exploitez.
- Sources de données : listez les données utilisées par votre solution.
- Exemples : rapports annuels, Yahoo Finance, Bloomberg, bases de données publiques, etc.
- Autres outils: mentionnez tout logiciel, librairie ou framework utilisé.



5. Bénéfices pour les utilisateurs

- Décrivez de manière concise les avantages concrets pour un analyste financier.
- Concentrez-vous sur les impacts mesurables ou perceptibles pour montrer la valeur ajoutée. Exemples: gain de temps, amélioration de la précision, prise de décision plus rapide, ou automatisation de tâches répétitives, etc.
- Conseil clé : mettez-vous à la place des entreprises présentes et pensez à ce qui retiendrait leur attention. La composition du jury est disponible sur le [site du Datathon](#).

6. Contacts

- Équipe : indiquez le numéro d'équipe ainsi que les noms des membres et profils LinkedIn.



4. CRITÈRES D'ÉVALUATION

Intelligence financière : pertinence et explicabilité (40%)

- Pertinence économique et financière : Les résultats et recommandations doivent être en lien direct avec la réalité économique et financière, et apporter une vraie valeur pour aider à la décision.
- Clarté et explicabilité : L'outil doit être capable d'expliquer clairement les choix effectués, ainsi que leurs impacts sur les décisions financières.
- Transparence et interprétabilité : Les modèles doivent être compréhensibles, avec des logiques claires et justifiées, pour inspirer confiance à l'utilisateur.
- Adaptabilité : Votre solution doit fonctionner avec n'importe quelle directive ou texte de loi fourni en entrée.

Utilisation des technologies et services AWS (25%)

- Qualité de l'architecture technique : Pertinence dans le choix des composants, modularité de la conception et robustesse de l'architecture.
- Exploitation des services AWS : Bonne utilisation des services proposés par AWS et des outils associés.
- Intégration et optimisation : Pertinence de l'intégration technologique avec d'autres environnements et optimisation de l'utilisation des ressources.
- Robustesse et scalabilité : Capacité de la solution à monter en charge et à fonctionner de manière fiable dans différents contextes d'utilisation.



Expérience utilisateur et fonctionnement de la solution (10%)

- Fluidité et fiabilité : Rapidité de l'application, stabilité et réactivité lors des différentes étapes d'utilisation.
- Lisibilité et visualisation : Qualité des visualisations proposées (graphiques, tableaux, schémas) et capacité à mettre en évidence les impacts financiers.
- Cohérence de l'expérience : Parcours utilisateur homogène, intuitif et adapté, de la première prise en main jusqu'à l'exploitation des résultats.

Présentation, storytelling et mise en avant de l'outil (25%)

- Clarté et impact de la présentation : Restitution claire et percutante des résultats, avec un fil conducteur évident.
- Storytelling et persuasion : Capacité à construire un récit convaincant qui met en valeur l'usage de l'outil et son intérêt pour l'utilisateur final.
- Différenciation et valeur ajoutée : Mise en avant des points forts et des éléments distinctifs qui rendent l'outil supérieur à d'autres solutions.
- Professionnalisme de la restitution : Respect des formats attendus, rigueur de la présentation et qualité visuelle.

Que les meilleurs gagnent

-L'équipe PolyFinances