



POLYFINANCES

DATATHON

2022

L'impact des banques centrales

MISE EN CONTEXTE

Les anticipations des politiques monétaires des banques centrales ont un impact important sur les marchés et sur l'évolution des devises. Sachant cela, un gestionnaire de fonds veut développer une stratégie d'investissement en plaçant son argent en CAD ou en USD sur une période donnée.

Ressources importantes

- Termes intéressants à explorer :
 - Fed jumping date
 - Sentometrics
 - Analyse textuelle
 - Huggingface
- Regardez la littérature sur le sujet !

Remise

- **Lundi, 7 nov. 12h** : NoteBook(s) pour entraîner et tester vos modèles avec vos résultats.
- **Judi, 10 nov. 23h** : Votre présentation PowerPoint.

DÉFI

Votre rôle comme expert en science des données est de concevoir un algorithme permettant de **prédire les fluctuations ou les valeurs** du taux de change CAD/USD sur une **période de deux semaines** à partir des données suivantes :

- Les communiqués de presse, les discours, les annonces et les transcriptions de la banque centrale des États-Unis.
- Les rapports de politique monétaire du Canada.
- Les taux directeurs.
- Autres données portant sur l'état économique de ces deux pays.

Après avoir réalisé une telle tâche, vous devez présenter votre solution au gestionnaire de fonds et celui-ci choisira la solution la plus prometteuse. Cela se fera sous forme de présentation devant jury.

PRÉSENTATION DEVANT JURY

1

Montrer l'impact des textes des banques centrales sur le taux de change USD/CAD.

Idées: Visualisation de données, statistique observée, lien macro-économique

2

Expliquer votre modèle de prédiction à partir des textes de la Fed ou d'autres données.

Idées: Modèles statistiques pour prédire le sentiment d'un texte, modèles d'apprentissage profond pour prédire l'impact.

3

Expliquer la démarche que vous avez utilisée pour réaliser un tel projet.

Idées: Partagez votre approche et ce que vous avez appris!

PRÉSENTÉ PAR



**BANQUE
NATIONALE**



FINANCE MONTRÉAL



CFA MONTRÉAL



S.R.A.I.

DONNÉES DISPONIBLES

1

Données de la Banque du Canada

Les responsabilités de la Banque du Canada sont axées sur les objectifs suivants : un taux d'inflation bas, stable et prévisible ; une monnaie sûre et sécuritaire ; un système financier stable et efficace au Canada et à l'échelle internationale. En pratique, elle a une définition interne plus étroite et plus précise de ce mandat : maintenir le taux d'inflation (mesuré par l'indice des prix à la consommation) entre 1 % et 3 %.

L'outil le plus puissant dont dispose la Banque du Canada pour atteindre cet objectif est sa capacité à fixer le taux d'intérêt de l'argent emprunté. En raison de l'importance des échanges commerciaux entre le Canada et les États-Unis, les ajustements spécifiques des taux d'intérêt sont souvent influencés par ceux des États-Unis à ce moment-là

→ Jeu de données venant de cette source:

- Taux de change USD - CAD [de 2017 à aujourd'hui] ← **Valeur/direction à prédire !**
- Taux d'intérêt auquel les banques commerciales appliquent à leurs clients les plus solvables. [de 2012 à aujourd'hui]
- Textes extraits des rapports des politiques monétaires de la Banque du Canada [2016 à aujourd'hui]

2

Données du comité fédéral de l'open market (FOMC) des U.S.

Aux États-Unis, le Federal Open Market Committee (FOMC), un comité du Federal Reserve System (la Fed), est chargé par la loi américaine de superviser les opérations des marchés de la nation. Ce comité de la Réserve Fédérale prend des décisions clés concernant les taux d'intérêt et la croissance de la masse monétaire des États-Unis.

La fourchette cible des fonds fédéraux est déterminée par une réunion des membres du Comité fédéral de l'open market (FOMC) qui se tient normalement huit fois par an à environ sept semaines d'intervalle. Le comité peut également tenir des réunions supplémentaires et mettre en œuvre des modifications du taux cible en dehors de son calendrier normal.

→ Jeu de données

- Tous les textes de la FOMC [de 1980 à aujourd'hui]
- Taux d'intérêt auquel les banques commerciales appliquent à leurs clients les plus solvables [de 2012 à aujourd'hui]

3

Autres données

→ Jeu de données

- Dictionnaire principal Loughran-McDonald avec listes de mots de sentiments dans l'optique de la finance (Libre à vous d'utiliser d'autres dictionnaires et même de créer le vôtre!)

Pour ce mandat, vous avez l'option de compléter ces jeux de données par toutes données que vous jugez pertinentes pour réaliser la tâche. En cas d'ajout de données, il sera important que vous démontriez au jury que les sources de vos données sont fiables et que vous les convainquiez que celles-ci pourront être utilisées dans le futur lors de besoin d'une prise de décision.



VOTRE MODÈLE

Comme vous avez plus le remarquer, le cas du Datathon de cette année est complexe. Afin de donner le plus de chance possible d'apprendre aux équipes, nous allons séparer celui-ci en trois étapes distincte.

1. Utiliser les **textes de la FED américaine pour sortir un sentiment sur le marché**. (40% des points)
 - À vous de choisir la meilleure façon d'extraire ce sentiment.
2. Refaire une **analyse similaire**, mais sur les textes de la **Banque du Canada** cette fois. (20% des points)
3. Utiliser ces sentiments (et autres métriques au choix) pour effectuer une **prédiction sur le taux de conversion** des devises des deux pays sur **un intervalle de deux semaines**. (40% des points)
 - Prédire le taux de change & si possible un intervalle d'incertitude ($X + 14$ jours)
 - 2022-05-02 (Predire: 2022-05-16), 2022-05-16 (Predire: 2022-05-30), ...
 - Sortez la métrique d'erreur MSE (erreur quadratique moyenne), le biais, et le drawdown de vos prédictions.

MÉTRIQUES D'ÉVALUATION

Pour la présentation

20% Qualité de la présentation et d'expression

60% Contenu de la présentation

- Votre approche au problème
 - Résumé du problème
 - Stratégie employée
 - Éléments externes que vous avez ajoutés
- Présentation des modèles
 - Explication des modèles
 - Validation effectuée
- Présentation des résultats
 - Prédiction
 - Limite des modèles
 - Ce que vous auriez aimé explorer
- Ce que vous avez appris

Pour les résultats

20% Afin de pouvoir comparer les différentes approches entre les équipes, vous allez devoir présenter les résultats de sentiments et de prévision **«out of sample» des 6 derniers mois**.

Cela signifie:

- vous pouvez utiliser toutes les données avant mai 2022 pour concevoir/entraîner/valider vos modèles.
- les résultats doivent provenir seulement des 6 derniers mois, mai 2022 à octobre 2022.

PRIX

1ÈRE PLACE

→ 2000\$

2ÈME PLACE

→ 1000\$

3ÈME PLACE

→ 500\$

PRIX DE PARTICIPATION

- Meilleure visualisation
- Meilleure vulgarisation
- Mini-jeux



HORAIRE

SAMEDI 5 NOVEMBRE - ATRIUM DU PAVILLON J.-A BOMBARDIER

- 9h00 - Accueil des participants avec viennoiseries et café
- 9h30 - Présentation de la compétition par PolyFinances
- 10h00 - Cours sur le NLP par SR.AI
- 10h30 - Pause
- 10h45 - Cours sur la finance par Banque Nationale
- 11h30 - Rencontre avec les commanditaires
- ~~~ Le reste de la journée se fera en virtuel
- Midi - Début du 48h de compétition + dévoilement des données
- 18h00 - Mini jeu #1
- 22h00 - Mini jeu #2

DIMANCHE 6 NOVEMBRE - VIRTUEL

- 10h00 - Période de questions avec Banque Nationale
- 14h00 - Vote de la meilleure visualisation de données + Prix !
- 18h00 - Mini jeu #3

LUNDI 7 NOVEMBRE - VIRTUEL

- Midi - Fin de la compétition + remise des projets

VENDREDI 11 NOVEMBRE - POLYTECHNIQUE MONTREAL

- 13h30 - Début des présentations devant jury à la Galerie Rolland
 - Horaire détaillé à venir
 - Remise des PowerPoint jeudi soir au plus tard
- 16h30 - Fin des présentations
- ~~~
- 17h30 - Soirée réseautage à l'Atrium du pavillon Lassonde !
 - Remise des prix
 - 10+ entreprises seront présentes
 - Des boissons seront offertes
- 19h30 - Fin de la soirée réseautage