

SEPTEMBRE 2023

FIAM

XPRESS

Faire le pont entre
la science des
données, IA/ML
et la gestion
des finances/actifs

LES GRANDS MODÈLES DE LANGAGE «EX. GPT» PEUVENT-ILS TRANSFORMER L'INDUSTRIE FINANCIÈRE ?

Évènement tenu dans les locaux de CDPQ le 24 août 2023

RAPPORT

de l'évènement « Bloomberg-CDPQ »
organisé par FIAM



De gauche à droite: Dr. Ranjan Bhaduri, David Ouellet, Jean-François Bérubé, Claude Perron, Dr. Russ Goyenko, Shawn Edwards et Dr. Hossein Kazemi.

Remerciements

Ce rapport a été rendu possible grâce à la contribution de plusieurs participants, Foulémata Tara Tiera, analyste à CDPQ, notamment, le rapport détaillé de Nicolas Charlton, directeur des stratégies quantitatives à la BNC.

Partenaires de l'évènement





SHAWN
EDWARDS

INTERVENTION DE SHAWN EDWARDS, CTO DE BLOOMBERG

POINTS SAILLANTS

Shawn Edwards a offert une étude détaillée des modèles de langage avancés, mettant en lumière leur évolution, leur connaissance globale et leurs compétences techniques. Il a retracé l'expérience de Bloomberg avec ces modèles, notamment la création et l'efficacité de BloombergGPT pour des tâches spécifiques à leur secteur. Edwards a mis en avant l'importance d'une mise en œuvre agile et les dangers potentiels liés à ces modèles.

Les Grands Modèles de Langage (LLM) : Une Révolution dans l'Industrie Financière

Les LLM sont en train de redéfinir la manière dont nous interagissons avec la technologie, en particulier dans des secteurs exigeants comme l'industrie financière. Aujourd'hui, je vais vous donner un aperçu de leurs capacités, de leurs applications et de leur potentiel.

Comprendre les LLM :

- **Nature Probabiliste** : Au cœur, les LLM sont des modèles probabilistes, prédisant le mot suivant dans une séquence.
- **Comportement Émergent** : Ce qui distingue les LLM, c'est leur capacité à comprendre le langage à un niveau avancé, démontrant des compétences de raisonnement qui étaient auparavant inaccessibles aux machines.
- **Réserve de Connaissances** : Imaginez avoir à votre disposition une bibliothèque dépassant la connaissance d'un individu moyen. C'est ce que les LLM offrent, une connaissance encyclopédique à portée de main.

Applications Techniques :

- **Assistant de Codage** (GitHub Copilot) : Pensez à un assistant qui, non seulement comprend votre demande, mais génère également du code pour vous. Les études montrent que cela peut booster la productivité des développeurs de 30 à 40 %.
- **La Distinction des LLM** :
 - **Polyvalence** : Les LLM, à la différence des anciens modèles d'IA, sont intrinsèquement polyvalents. Ils peuvent être adaptés à une multitude de tâches sans nécessiter de formation spécialisée.
 - **Étendue de la Connaissance** : Ils ne sont pas seulement vastes, ils sont profonds, offrant une compréhension détaillée de nombreux sujets.

- o **Accessibilité** : L'interaction avec les LLM se fait en langage naturel, pas en code. Cela ouvre la porte à une démocratisation de leur utilisation. Les experts financiers, par exemple, peuvent désormais concevoir des systèmes sans dépendre entièrement des développeurs.

Une Rétrospective sur les Grands Modèles de Langage (LLM)

L'évolution des LLM : Historiquement, les modèles de langage visaient à prédire le mot suivant dans une séquence. Cependant, ce qui fait la singularité des LLM modernes est leur aptitude à comprendre et à intégrer l'ensemble du contexte lors de cette prédiction. Ils ne se contentent pas de suivre une séquence linéaire, mais évaluent l'ensemble du paysage sémantique pour déterminer la probabilité du mot suivant.

L'Impact de l'Échelle et du Développement : L'ascension fulgurante de l'IA n'aurait pas été possible sans l'expansion phénoménale de la puissance de calcul. Les théories sur lesquelles reposent les LLM ne sont pas nouvelles, mais c'est la combinaison de ces théories avec des ressources informatiques sans précédent qui a rendu possible l'apprentissage de modèles d'une telle envergure. Bien que l'entraînement de ces géants puisse prendre des mois sur des infrastructures cloud, le résultat est un modèle capable de tâches auparavant inimaginables.

Applications et Cas d'Utilisation :

- **Traitement de Documents** : Les LLM ont révolutionné le traitement de l'information. Qu'il s'agisse de synthétiser des informations, d'extraire des données clés, de comprendre des documents complexes ou de générer des textes, leur capacité à traiter le langage est sans égale.
- **Génération de Code** : La capacité des LLM à traduire le langage naturel en code est un tournant. Cela réduit la barrière entre l'idéation et la création, permettant à ceux qui ne sont pas familiers avec la programmation de contribuer au développement d'outils.
- **Intégration de l'API** : Chez Bloomberg, nous explorons activement comment les LLM peuvent être formés pour interagir avec nos API. L'idée est de permettre une interaction plus fluide entre nos systèmes et ces modèles, ouvrant la voie à des applications encore plus innovantes.

Le Parcours de Bloomberg avec les LLMs (BloombergGPT)

Développement : Bloomberg a consacré deux ans à la conception et au développement de BloombergGPT (BBGPT). L'objectif principal était d'explorer si l'intégration de données spécifiques à un domaine pouvait optimiser les performances du modèle GPT dans ce secteur précis.

Sources de Données pour l'Entraînement : Pour former BloombergGPT, nous avons utilisé une combinaison judicieuse de données publiques et de données financières spécifiques. Ces données financières sont le fruit des efforts considérables de Bloomberg en matière de création, de collecte, de conservation et de recherche de données.

Résultats Observés : BloombergGPT a montré une compétence impressionnante dans les tâches générales. Cependant, là où il a vraiment brillé, c'est dans les tâches spécifiques à un domaine. Ces résultats confirment l'hypothèse selon laquelle les organisations qui possèdent des données de haute qualité et spécifiques à un domaine ont un avantage compétitif notable.

L'Importance de la Formation Spécifique à un Domaine : Au début, nous étions sceptiques quant à l'impact de la formation sur des données spécifiques à un domaine. Cependant, nos expériences avec BloombergGPT ont clairement démontré que cette approche ciblée peut produire des améliorations substantielles dans les performances du modèle.

L'Ingénierie des Invites (*Prompt*) chez Bloomberg

L'Importance de l'Ingénierie des Invites (*Prompt*) : Chez Bloomberg, nous reconnaissons que la manière dont on interagit avec les modèles comme GPT peut grandement influencer la qualité des réponses obtenues. C'est pourquoi nous investissons activement dans l'ingénierie des invites. Notre objectif est d'affiner ces invites pour permettre aux utilisateurs d'extraire des informations plus précises et pertinentes des modèles GPT.

Meilleures Pratiques pour l'Ingénierie des Invites :

- 1. Fournir un Contexte :** Les modèles GPT répondent mieux lorsqu'ils ont une compréhension claire de ce qui est demandé. En fournissant un contexte, vous donnez au modèle une base solide pour générer une réponse.
- 2. Utiliser des Exemples :** Les exemples aident à guider le modèle vers le type de réponse que vous attendez. C'est comme montrer un modèle de ce que vous voulez obtenir.
- 3. Stratégie d'Approfondissement :** Commencez par une invite générale pour établir le sujet, puis approfondissez avec des détails, du contexte et des exemples. Par exemple, au lieu de simplement demander "Qu'est-ce que le plafond de la dette?", commencez par une invite comme "Recherchez des informations sur le plafond de la dette" et suivez avec des questions plus spécifiques pour obtenir des détails ou des clarifications.

Fiabilité et Meilleures Pratiques avec les LLM

Méthodes non fiables : Lorsqu'on utilise des documents longs sans fournir de contexte clair, les LLM peuvent facilement dériver et produire des résultats qui ne sont pas pertinents ou inexacts.

Méthodes fiables : Pour obtenir des réponses précises, il est préférable d'utiliser des documents d'entrée courts et bien structurés, accompagnés d'un contexte clair. Cela guide le modèle et augmente la probabilité d'obtenir une réponse pertinente.

Projets Futurs de Bloomberg avec les LLM

Intégration Opérationnelle : Bloomberg est actuellement en phase d'exploration sur la manière d'intégrer BloombergGPT dans ses applications basées sur l'IA.

Cas d'utilisation potentiels :

- 1. Conversion de Textes :** Transformer le langage naturel en langage de requête Bloomberg (BQL) pour faciliter les requêtes.
- 2. Interface Linguistique :** Imaginer une interface pour le terminal Bloomberg où les utilisateurs peuvent interagir en langage naturel.
- 3. Copilote pour Analystes :** Un assistant qui peut résumer, détecter des tendances, intégrer des recherches et bien plus encore, le tout pour aider les analystes dans leur travail quotidien.

Risques Associés aux LLM:

- **Droits d'Utilisation :** Qui détient les droits sur le contenu généré par le modèle?
- **Confidentialité :** Comment s'assurer que les données sensibles ne sont pas exposées?
- **Exigences d'Archivage :** Comment stocker et archiver les résultats générés?
- **Provenance du Contenu :** Comment déterminer la source du contenu généré?

Nouveaux Risques : Avec l'avènement des LLM, de nouveaux défis se présentent. L'un des plus préoccupants est le risque que le modèle produise des informations qui sont non seulement inexacts, mais aussi potentiellement trompeuses.



**DR. RUSS
GOYENKO**

PRÉSENTATION PAR LE DR. RUSS GOYENKO (MCGILL)

POINTS SAILLANTS

Russ Goyenko a exposé une étude démontrant la capacité de l'IA à décrypter avec précision les intentions des dirigeants d'entreprise. Cette étude a utilisé l'IA pour anticiper les surprises financières en se basant sur les commentaires de gestion dans les rapports d'entreprise.

Titre : “Les grands modèles de langage peuvent-ils fournir des prévisions plus précises aux analystes?”

Introduction : Basée sur l'article de recherche intitulé “L'IA peut-elle lire dans l'esprit des dirigeants d'entreprise?”, cette présentation explore la capacité des LLM à prédire les surprises en matière de bénéfices en analysant des documents financiers. (source: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4493166)

Points Clés :

- 1. Les LLM comme Outils Prédicatifs :** Les LLM, après une formation approfondie sur les documents 10Q et 10K et une analyse exhaustive des archives EDGAR, ont démontré une capacité impressionnante à prédire avec précision les surprises en matière de bénéfices, surpassant les prévisions consensuelles des analystes traditionnels.
- 2. Validation Économique :** En utilisant les prédictions des LLM, une stratégie d'investissement Long-Short a été élaborée. Cette stratégie, qui consiste à acheter les actions les plus performantes et à vendre à découvert les moins performantes, a généré un alpha annuel ajusté au risque du marché de 6 %.
- 3. Mécanisme Sous-Jacent :** La clé de cette réussite réside dans les informations contenues dans les documents 10-Q et 10-K. Ces documents regorgent de données que les humains ne peuvent pas facilement traiter. Les LLM, en revanche, peuvent analyser ces vastes quantités d'informations, captant les nuances et affinant leurs prévisions avec chaque nouvelle donnée.
- 4. Conclusion :** L'ère actuelle exige une intégration poussée des technologies avancées dans le processus décisionnel financier. Les LLM, avec leur capacité à analyser et à prédire à partir d'énormes volumes de données, représentent l'avenir de la prise de décision financière précise. Pour les investisseurs et les analystes, embrasser ces technologies est non seulement judicieux, mais essentiel pour rester compétitif.

Présentation du rapport

Russ Goyenko a présenté les conclusions d'un rapport de recherche sur la capacité de l'IA à "lire dans les pensées" des dirigeants d'entreprise. L'essentiel de la recherche suggère que l'IA peut effectivement interpréter les intentions des dirigeants d'entreprise avec un certain degré de certitude investissable.

Principales informations

Prédire les fondamentaux à l'origine des rendements : La recherche visait à prédire les fondamentaux connus pour influencer les rendements. Pour ce faire, elle a examiné les résultats trimestriels inattendus par rapport aux attentes du marché, en particulier à partir des formulaires 10-Q et 10-K disponibles sur EDGAR. La section "Management Discussion & Analysis" (MD&A) des rapports trimestriels et annuels a été la principale source de données.

Volume de données :

La section moyenne du rapport de gestion contient environ 10 000 caractères. Compte tenu du grand nombre d'entreprises et de leurs rapports annuels, il n'est pas possible pour les humains de lire chaque rapport de gestion. Il est intéressant de noter que les machines représentent déjà 75 % des téléchargements EDGAR.

Approches d'apprentissage automatique :

Les modèles traditionnels d'apprentissage automatique s'appuient sur un dictionnaire de mots, qu'ils comparent aux mots contenus dans les sections du rapport de gestion. Cependant, la direction de l'entreprise a commencé à manipuler ce système.

La recherche a permis d'introduire une nouvelle méthode de modèle de langage étendu (LLM) capable de traiter des documents de n'importe quelle longueur. Ces modèles d'IA ont été entraînés à prédire les prochaines surprises en matière de bénéfices 10-Q. En utilisant une stratégie Long-Short (L-S) pour les 20% supérieurs et inférieurs des surprises prédites, le modèle a surperformé le S&P 500 de 6,74% par an (6,01% sur une base ajustée au risque).

Test du modèle BERT :

Un modèle générique de BERT et un modèle de BERT affiné ont été testés pour leur capacité à prédire les surprises liées aux bénéfices. Le BERT affiné, utilisé dans une stratégie à long terme, a surpassé le S&P 500 par un facteur de 4 entre 2004 et 2022.

La longueur du rapport de gestion en tant qu'indicateur prévisionnel :
La longueur de la section MD&A s'est avérée être un indicateur significatif, presque aussi efficace qu'un modèle BERT non affiné.

Les rapports excessivement longs ont eu tendance à sous-performer. À l'inverse, les rapports trop courts ont souvent surperformé. La recherche a atteint un taux de réussite de 60 % dans la prédiction de la direction des surprises en matière de bénéfices.

Informations complémentaires - échantillon de formation :

L'échantillon de formation du modèle a été mis à jour chaque année. Au début de chaque année, le modèle a été réentraîné à l'aide des données les plus récentes et ses poids ont été fixés pour l'année en question. Ce processus a été répété au début des années suivantes.

POINTS CLÉS DES DIRIGEANTS DE LA CDPQ ET DES PSP



**JEAN-FRANÇOIS
BÉRUBÉ**

Jean-François Bérubé de la CDPQ, un groupe d'investissement mondial actif sur les principaux marchés financiers, a introduit le sujet en soulignant le potentiel de transformation de l'IA générative (GenAI), citant une comparaison souvent répétée avec les possibilités permises par l'invention de l'ordinateur personnel. Il a présenté un bref historique de la ML/AI, retraçant son évolution depuis les modèles simples jusqu'aux systèmes avancés tels que ChatGPT.

Les propositions de valeur de GenAI comprennent l'identification des opportunités d'investissement, l'analyse des risques du portefeuille et l'amélioration de l'efficacité opérationnelle.

- **Offensive** : aider à trouver des opportunités d'investissement et des tendances émergentes sur le marché.
- **Défensive** : aider à analyser l'exposition d'un portefeuille aux risques liés à l'innovation.
- **Numérique** : Améliorer l'aspect opérationnel de l'entreprise. David (Ouellet) voit de nombreuses opportunités sur ce front à court terme.



DAVID OUELLET

David Ouellet d'Investissements PSP, un acteur majeur du secteur privé et le plus grand propriétaire agricole au monde.

PSP a utilisé les innovations technologiques pour développer son propre PSPGPT, qui est maintenant utilisé quotidiennement par environ un tiers de son équipe. David mentionne l'utilisation d'un outil d'IA, le "ChatPSP", pendant les réunions pour améliorer les discussions en suscitant des questions et des idées. Il est intéressant de noter qu'ils utilisent le modèle pour se remettre en question et réduire le risque que des biais comportementaux se glissent dans leur décision finale.

Alors que les membres du conseil d'administration pourraient vouloir augmenter le budget des projets GenAI, David souligne l'importance de donner la priorité au stockage et à l'accessibilité des données dans le nuage, car ils sont essentiels à la réussite des futurs projets GenAI. Il a également mentionné que l'utilisation de services externes tels que BloombergGPT reste un moyen probable. PSP élabore actuellement de nombreuses preuves de concept pour le front office afin d'accroître l'efficacité de ses gestionnaires de projets.



**DR. HOSSEIN
KAZEMI**

POINTS SAILLANTS

Le Dr. Kazemi a insisté sur la nécessité pour les experts financiers de maîtriser GenAI. Il a évoqué les initiatives de la CAIA pour lancer une nouvelle certification professionnelle, le FDP "Financial Data Professional", déjà reconnue dans 40 pays.



**DR. RANJAN
BHADURI**

DISCUSSION EN PANEL

Ranjan Bhaduri a animé les discussions et les questions sur ...

POINTS SAILLANTS

Lors de la discussion sur l'IA générative en finance, les spécialistes ont mis en avant les enjeux éthiques et ESG de l'IA, soulignant l'importance d'une utilisation responsable et d'une intégration mesurée. L'impact croissant de l'IA dans la décision financière et l'éducation a été abordé, tout comme les obstacles liés à la disponibilité limitée des données.

ÉTHIQUE DE L'IA ET PRÉOCCUPATIONS ESG

David Ouellet: Met l'accent sur une approche équilibrée, considérant l'importance de peser à la fois les risques et les avantages potentiels lorsqu'il s'agit de décider de l'utilisation de cette technologie.

Shawn Edwards: Il souligne les dangers de s'appuyer sur des modèles statistiques construits à partir de données biaisées, qui peuvent involontairement perpétuer des préjugés. Il s'inquiète de l'utilisation malveillante des modèles d'IA, notamment pour influencer les élections, créer des courriels d'hameçonnage indétectables et produire des "deepfakes". Il souligne l'importance de trouver des moyens d'authentifier les résultats de l'IA par rapport à ceux de l'homme.

Dr. Hossein Kazemi: souligne les risques posés par les utilisateurs non avertis qui ne comprennent pas les complexités de l'IA, en établissant un parallèle avec les investisseurs mal informés pendant la crise financière.

SYSTÈMES D'AUTO-APPRENTISSAGE COMME ALPHA GO ET LLMS

Dr. Bhaduri: AlphaGo a appris le jeu de Go en examinant des parties jouées par des champions humains, et est devenu le célèbre champion du monde. Cependant, AlphaGo Zero, qui était un système d'auto-apprentissage, a battu AlphaGo 100-0. AlphaGo Zero avait des mouvements créatifs qui n'étaient pas du tout conventionnels. Avec les LLM, nous essayons d'apprendre à partir de modèles humains... est-ce que cela limite le potentiel d'amélioration ?

Dr. Russ Goyenko: Il préconise la conception de "gestionnaires de portefeuille" d'IA qui surpassent constamment leurs pairs dans tous les scénarios. Il laisse entendre qu'il a mené des recherches spécifiques sur ce sujet.

LE RÔLE DE L'IA DANS LES COMITÉS D'INVESTISSEMENT

Jean-François Bérubé: note que certains modèles d'IA sont déjà utilisés pour améliorer les processus de prise de décision.

David Ouellet: mentionne l'utilisation d'un outil d'IA, "ChatPSP", pendant les réunions pour améliorer les discussions en suscitant des questions et des idées. Il est intéressant de noter qu'ils utilisent le modèle pour se remettre en question et réduire le risque que des biais comportementaux se glissent dans leurs décisions finales.

L'IA DANS L'ÉDUCATION

Dr. Hossein Kazemi : il évoque l'utilisation d'un outil qui estime la probabilité que le contenu soit généré par l'IA. Il mentionne également une approche innovante dans laquelle les étudiants transmettent à l'IA les informations contenues dans les manuels scolaires et demandent ensuite à l'IA de leur poser des questions.

L'IA DANS LA RECHERCHE EN INVESTISSEMENT

Jean-François Bérubé: Il fait part du développement d'un outil destiné à aider les analystes en recherche d'actions, en particulier à examiner et à résumer les transcriptions des appels à bénéfices en grandes quantités et à un rythme rapide, ce qui nécessiterait des ressources considérables s'il était effectué par des employés.

Shawn Edwards: Parle de l'identification des tendances en matière de rapports de gestion dans différentes entreprises.

LES DÉFIS POSÉS PAR LES CHAMPS DE DONNÉES ÉPARSES

Dr. Russ Goyenko (McGill): souligne les limites des grands modèles lorsqu'ils travaillent avec des données éparées, suggérant que se fier uniquement aux données mensuelles sur les prix nécessiterait plus d'un millénaire pour un entraînement efficace du modèle.

Appliquez-vous le test de Turing ? ...

RÉSUMÉ DES QUESTIONS OUVERTES

L'avenir des emplois dans la finance

Shawn Edwards Edwards explique qu'une grande partie de leurs articles étaient déjà générés par ordinateur, ce qui permettait aux journalistes d'approfondir les sujets. Il mentionne également l'évolution de leurs équipes de données globales, qui sont passées d'employés de bureau à des formateurs de modèles. Il note toutefois que les codeurs qui utilisent des outils tels que Copilot ont tendance à produire un code moins sûr et plus sujet aux erreurs.

Dr. Hossein Kazemi estime qu'il est prématuré de prédire les changements exacts, mais reconnaît le caractère inévitable de l'évolution du paysage professionnel.

L'utilisation de la GenAI pour l'efficacité opérationnelle

Jean-François Bérubé déclare qu'ils adoptent GenAI pour soutenir l'efficacité opérationnelle au Front Office.

Alimenter les LLM en documents volumineux

Shawn Edwards suggère de diviser les gros documents en plus petits morceaux et de les introduire dans le modèle de manière itérative, puis de combiner les résultats.

Les risques liés à l'utilisation de l'IA par des investisseurs non avertis

Dr. Hossein Kazemi (CAIA) affirme qu'il est très probable qu'une crise financière soit provoquée par l'utilisation d'outils d'IA par des investisseurs non avertis. Il envisage un scénario dans lequel quelques services d'investissement à base d'IA dominants pourraient, par inadvertance, orienter les investisseurs vers les outils d'IA. services d'investissement en IA pourraient par inadvertance orienter un grand nombre d'investisseurs dans la même direction, ce qui amplifierait les risques.