

Étude de cas

MIROCLUS

SCIENCE SIMPLIFIED

Miroculus

Machine Learning et recherche médicale accélérée grâce à la technologie des graphes

SECTEUR

Sciences de la vie

CAS D'USAGE

Graphe de connaissances

OBJECTIF

Trouver le bon biomarqueur micro-ARN pour dépister le cancer de l'estomac par test sanguin.

DÉFI

Analyser les résultats de la recherche médicale volumineuse sur les micro-ARN.

SOLUTION

Créer un graphe de connaissances des dernières recherches sur les micro-ARN.

RÉSULTATS

- Compréhension approfondie et accélérée de la recherche médicale en cours
- Identification plus rapide des micro-ARN indiquant un cancer de l'estomac

Miroculus a créé une plateforme qui détecte les micro-ARN dans le sang. Pour parcourir le recueil toujours plus important de publications médicales pertinentes sur le sujet, Miroculus a conçu un pipeline d'apprentissage machine qui saisit la recherche sur les micro-ARN, les gènes et les maladies dans un graphe de connaissances stocké dans Neo4j.*

L'entreprise

Miroculus, une entreprise de diagnostics moléculaires basée à San Francisco, a développé une plateforme numérique dédiée à la microfluidique qui permet aux scientifiques et aux professionnels de la santé de détecter des maladies en utilisant les micro-ARN. La plateforme de diagnostics traite différents types de tests. Elle est flexible et facile à utiliser.

Le défi

Pour la plupart des types de cancer, un diagnostic précoce améliore le taux de survie des patients. Le problème est que beaucoup des tests utilisés pour diagnostiquer le cancer sont invasifs et nécessitent un équipement spécial.

Par exemple, le cancer de l'estomac est toujours diagnostiqué avec une endoscopie, une technologie qui date de 1965. Par ailleurs, les endoscopies ne sont pas un outil de dépistage efficace : seuls 2 % des patients dépistés ont un cancer de l'estomac. Une méthode de diagnostic plus simple pourrait éliminer les procédures inutiles et améliorer le dépistage.

Miroculus a entrevu la promesse que pouvaient apporter les micro-ARN dans le dépistage du cancer, mais on croyait alors les micro-ARN enfermés dans des cellules. En 2008, un des conseillers de l'entreprise a découvert que des micro-ARN circulaient. Lorsqu'il y a un problème dans les cellules au niveau du tissu, elles se séparent et libèrent des micro-ARN dans le système sanguin.

Aujourd'hui, détecter les micro-ARN nécessite encore des scientifiques hautement qualifiés, des réactifs onéreux, des machines et des protocoles extrêmement complexes mais Miroculus y a vu le potentiel de changer la face du dépistage de micro-ARN.

Mener une étude pour trouver un biomarqueur de micro-ARN pour le cancer de l'estomac signifiait se tenir au courant d'une percée dans la recherche médicale concernée. Les méthodes de recherche principales ont impliqué la recherche et la sélection d'articles pertinents, la tentative de les récupérer et de les assimiler.

Avec l'augmentation de la recherche sur les micro-ARN, intégrer toutes les informations pertinentes prendrait énormément de temps. Miroculus avait donc besoin d'un moyen d'accélérer ce processus et de connecter les scientifiques directement à la recherche concernée.

La solution

Trouver un biomarqueur de micro-ARN pour le cancer de l'estomac est un domaine de recherche incontournable. Parmi le million de personnes diagnostiquées d'un cancer de l'estomac, il y en a 80 % qui survivent moins de 18 mois.

Étude de cas



« Afin de donner un sens à toute l'information nouvellement disponible sur les micro-ARN, nous avons stocké cet important volume de données dans une base de données de graphes consultable. »

– Antonio Molins,
Vice Président Data Science
chez Miroculus

Leur travail a nécessité de rester au courant des dernières publications qui relient les gènes, les maladies et les micro-ARN. « Afin de donner un sens à toute l'information nouvellement disponible sur les micro-ARN, nous avons stocké ce gros volume de données dans une base de données de graphes consultable, » a déclaré Antonio Molins, Vice-Président Data Science chez Miroculus.

L'équipe Miroculus a réuni plus d'un milliard d'articles dans Hadoop. Puis, ils ont utilisé le traitement du langage naturel (TLN) pour extraire des phrases précises avec des mots-clés pour gène, maladie et micro-ARN. Déduire la relation entre les mots-clés a cependant nécessité une autre étape. L'équipe a développé un modèle d'apprentissage machine autonome pour classifier les relations qui sont ensuite stockées dans Neo4j.

« Nous croyons qu'il est important d'utiliser le bon outil pour le bon problème, » a affirmé Molins. « Les bases de données de graphes sont le bon outil si vous vous concentrez sur les relations. »

L'équipe a créé une visualisation interactive consultable par micro-ARN, gène et maladie. Le critère de recherche de l'utilisateur devient le nœud central de la visualisation. Il est entouré de nœuds qui relient des micro-ARN précis. Sélectionner un nœud en particulier fait remonter des articles qui les relient à la phrase spécifique citée et à un lien vers la publication.

Avec les dernières recherches médicales en main, Miroculus a conçu une étude qui suit les directives de la FDA (Food and Drug Administration, « Agence américaine des produits alimentaires et médicamenteux »). L'étude, conduite en collaboration avec le NIH (Instituts Américains de la Santé), l'Institut national du cancer et des experts au Chili, a porté sur 650 personnes éligibles pour une endoscopie afin de diagnostiquer le cancer de l'estomac.

Les résultats

L'équipe Miroculus a découvert plus rapidement des liens pertinents entre les micro-ARN et le cancer de l'estomac avec la recherche médicale connectée à Neo4j. « Cela nous a beaucoup aidé à trouver les micro-ARN spécifiques liés au cancer de l'estomac, » a estimé Molins.

L'équipe Miroculus a identifié le micro-ARN qui permet de dépister avec succès le cancer de l'estomac à l'aide d'un échantillon de sang. Leur plateforme propose une approche peu invasive qui permet de réserver l'endoscopie à ceux qui en ont vraiment besoin. « Le test fonctionne à différents stades de la maladie, aussi bien au début qu'à un stade ultérieur. C'est important car nous voulons diagnostiquer les personnes atteintes du cancer à un stade précoce, là où le pronostic est bien meilleur, » déclare Molins.

Miroculus a rendu public sa visualisation des liens entre les gènes, les maladies et les micro-ARN. La visualisation codée par couleur interroge la base de données Neo4j de façon interactive. « Les utilisateurs trouvent beaucoup plus de valeur dans le fait d'aller d'abord aux conclusions qu'ils veulent examiner, puis de récupérer l'article, plutôt que l'inverse. Cela semble plus efficace, » a affirmé Molins.

La sélection d'un élément récupère la phrase précise dans l'article qui comprend les éléments de recherche, ainsi que le lien vers l'article complet, ce qui permet de connecter les chercheurs aux études pertinentes.

« L'apprentissage machine et la technologie des graphes comme Neo4j ont permis d'accélérer la recherche scientifique dans le domaine médical, » a expliqué Jorge Soto, cofondateur et directeur technique chez Miroculus.

Neo4j est le leader de la technologie des bases de données de graphes. Avec le plus grand nombre de déploiements au monde, Neo4j aide des entreprises mondiales - comme Airbus, [Michelin](#), [NASA](#), [Crédit Agricole](#) et [Volvo Cars](#) - à prédire et identifier la façon dont les personnes, les processus, les lieux et les systèmes sont interconnectés. Grâce à cette approche par les relations, les [applications](#) mises au point en utilisant Neo4j relèvent les défis associés aux données connectées, tels que l'[analytique et l'intelligence artificielle](#), la [détection de fraude](#), les [recommandations](#) en temps réel et les [graphes de connaissance](#). Pour en savoir plus, merci de consulter [Neo4j.com](#) et [@Neo4jFr](#).

Des questions sur Neo4j ?

Contactez-nous :
info@neo4j.com
neo4j.com/contact-us