vinetur

Une étude révèle que les modèles d’apprentissage automatique n’identifient les cépages de Macédoine du Nord qu’avec une précision de 56 %

Les chercheurs ont testé 1 024 vins de plusieurs millésimes ; les modèles appliqués aux vins blancs n’ont atteint qu’environ 49 % de précision sur des échantillons inédits

mardi 17 novembre 2026



Une étude publiée dans la revue Beverages a révélé que les modèles d'apprentissage automatique ne parvenaient que dans une mesure limitée à identifier les cépages des vins de Macédoine du Nord à partir de mesures physico-chimiques courantes. Elle met en évidence l'écart entre des données de laboratoire prometteuses et une classification fiable en conditions réelles.

Les chercheurs ont analysé 1 024 vins issus de plusieurs millésimes et cherché à déterminer si des résultats d'analyses standard pouvaient servir à distinguer les cépages. Selon l'étude, les modèles ont obtenu des résultats modestes sur des échantillons jamais vus, avec une précision d'environ 56 % pour les vins rouges et de

49 % pour les vins blancs. La réduction du nombre de variables d'entrée au moyen de l'analyse en composantes principales (ACP) a détérioré les performances au lieu de les améliorer.

Ces résultats mettent en lumière un défi fondamental de l'analyse des vins : des données faciles à recueillir ne suffisent pas toujours à établir une classification solide. Les analyses physico-chimiques courantes peuvent renseigner sur les caractéristiques générales d'un vin, telles que sa composition et son style, mais l'étude laisse entendre qu'elles ne fournissent peut-être pas une empreinte assez distincte pour différencier systématiquement les cépages d'un millésime à l'autre.

L'enjeu dépasse le cadre de la recherche universitaire. Dans le secteur des boissons, les entreprises et les autorités de réglementation s'intéressent de plus en plus à l'intelligence artificielle pour faciliter les contrôles d'authenticité, le contrôle qualité et la vérification des produits. Ces nouveaux résultats incitent à ne pas surestimer les capacités de ces outils lorsqu'ils reposent principalement sur des variables de laboratoire standard. Pour les domaines viticoles, les importateurs et les autres acteurs de la filière, l'étude rappelle que les systèmes d'apprentissage automatique pourraient nécessiter une validation plus rigoureuse et des données plus riches avant d'être utilisés pour des décisions concernant l'étiquetage, la traçabilité ou la détection des fraudes.

Le fait que l'étude porte sur plusieurs millésimes est également important. Les variations d'un millésime à l'autre peuvent modifier le profil mesurable d'un vin en fonction des conditions météorologiques, des modalités de maturation et des choix de production. Un modèle qui semble précis sur un jeu de données restreint peut devenir moins fiable lorsqu'il est confronté à des vins produits d'autres années. Les résultats plus faibles obtenus lors des tests montrent qu'il reste difficile de généraliser d'un millésime à l'autre, même lorsque les vins proviennent du même pays et relèvent de la même catégorie d'analyse.

Ces résultats apportent également une nuance au débat plus large sur l'intelligence artificielle dans l'analyse des aliments et des boissons. L'apprentissage automatique peut détecter des schémas qui échappent à l'observation humaine, mais ses résultats

dépendent fortement de la nature et de la qualité des données utilisées. Dans le cas présent, les informations contenues dans les mesures physico-chimiques courantes semblent avoir été limitées. Les résultats moins bons après l'ACP suggèrent que la réduction du jeu de données à un plus petit nombre de variables composites a supprimé des informations dont les modèles avaient encore besoin.

L'étude ne signifie pas que l'apprentissage automatique n'a pas sa place dans la classification des vins. Elle indique plutôt que les seules données physico-chimiques courantes ne suffisent peut-être pas à identifier précisément les cépages dans un contexte exigeant et concret. Des méthodes d'analyse plus spécialisées, des jeux de données plus vastes ou des modèles conçus pour des tâches de classification plus ciblées pourraient encore améliorer les résultats. Toutefois, les éléments présentés ici montrent que les affirmations concernant l'authentification rapide par l'IA doivent être vérifiées avec soin avant que ces outils ne soient utilisés à des fins commerciales ou réglementaires.

Cette distinction est importante pour les producteurs et les acheteurs. Un système qui obtient des résultats raisonnables lors de l'entraînement, mais ne classe correctement qu'environ la moitié des échantillons testés, serait difficile à utiliser pour des décisions à forts enjeux. Dans le secteur du vin, où l'origine, le cépage et l'authenticité peuvent influer directement sur le prix et la confiance des consommateurs, un taux d'erreur de ce niveau pourrait entraîner des difficultés pratiques et juridiques s'il n'était pas compensé par des vérifications supplémentaires.

L'étude invite donc à une approche plus prudente de l'IA dans l'analyse des vins. Elle montre que l'apprentissage automatique peut repérer des schémas dans la composition chimique courante des vins, mais aussi que ses performances peuvent rapidement diminuer lorsqu'il doit classer de nouveaux échantillons issus de plusieurs millésimes. Pour le secteur des boissons, il ne s'agit pas tant de renoncer à l'IA que de définir des attentes réalistes et d'exiger une validation externe rigoureuse avant de considérer ces systèmes comme des outils fiables.