جامعة الملك سعود

عمادة الدراسات العليا

تأثير تطبيق الأسس الإحصائية الاحتمالية في بناء النماذج التوليدية الحديثة المستخدمة في تقييم أداء الشركات السعودية. " مُراجعة أدبيات"

The Impact of Applying Probabilistic Statistical Foundations in Building Modern Generative Models Used for Evaluating the Performance of Saudi Companies: A Literature Review

خطّة بحث مقدّمة استكمالًا لمتطلّبات الحصول على درجة الماجستير في الإحصاء

إعداد الطالب:

_____

(الرقم الجامعي يوضع بين القوسين)

إشراف:

_____

أستاذ ----بقسم الدراسات الإسلاميّة

الفصل الدراسي 2026

1445هـ

بسم الله الرحمن الرحيم

مقدمة

شهد العالم في السنوات الأخيرة تطورًا كبيرًا في مجال الذكاء الاصطناعي؛ حيث ظهرت نماذج توليدية حديثة قادرة على إنتاج محتوى جديد يشبه الإبداع البشري، وهذه النماذج انتقلت بسرعة من المختبرات البحثية إلى الاستخدامات اليومية في مختلف المجالات، مثل (الطب والهندسة والاقتصاد والتعليم والأمن السيبراني)، وهذا التطور جعل من مجال الذكاء الاصطناعي واحدًا من أكثر مجالات البحث العلمي تطورًا وتأثيرًا في الوقت الحالي(Goodfellow et al., 2016؛ Bommasani et al., 2021؛ OpenAI, 2023).

وعلى الرغم من أن هذه النماذج غالبًا ما تعتبر إنجازًا في علوم الحاسب والذكاء الاصطناعي؛ إلا أن أساسها العلمي يعتمد على مفاهيم إحصائية ومعرفة متقدمة، وعملية توليد البيانات هي في الأساس محاولة لنمذجة التوزيع الاحتمالي الحقيقي للبيانات، ثم استخدام هذا النموذج لإنشاء عينات جديدة تحافظ على الخصائص الإحصائية للبيانات الأصلية، ومن هنا يظهر دور الإحصاء بشكل واضح في بناء هذه النماذج، حيث تعتمد على نظريات الاحتمالات والاستدلال الإحصائي والتقدير الإحصائي وتحسين الدوال الهدف وقياس عدم اليقين (Murphy, 2023؛ Bishop & Bishop, 2024).

حيث تطورت القدرة الحاسوبية وتوافر البيانات الضخمة أدى إلى ظهور أجيال متعاقبة من النماذج التوليدية، بدءًا من النماذج الاحتمالية التقليدية، مرورًا بالشبكات التنافسية التوليدية والمشفرات التباينية، وصولًا إلى نماذج الانتشار والنماذج اللغوية الكبيرة، رغم اختلاف البنية الرياضية لهذه النماذج؛ فإنها تشترك في هدف إحصائي واحد، يتمثل في تعلم التوزيع الكامن للبيانات وإنتاج بيانات جديدة ذات خصائص مشابهة بدرجة عالية من الواقعية (Kingma &

Brown et al., ؛ Ho et al., 2020؛ Goodfellow et al., 2014؛Welling, 2014 .(2020

مع هذا التقدم السريع، ظهرت تحديات علمية مهمة تتعلق بموثوقية النماذج التوليدية ودقة مخرجاتها، ومن المهم تقييم أداء هذه النماذج لضمان تمثيل الواقع بدقة دون تحيز أو تشويه؛ لذلك أصبح تقييم الأداء أحد المحاور الرئيسة في الأبحاث الحديثة، ويتجاوز الحكم على جودة النموذج إلى تقييم جودة التوزيع الاحتمالي وقياس درجة التشابه الإحصائي بين البيانات الحقيقية والمولدة (Bishop & Bishop, 2024؛Borji, 2019).

مشكلة البحث

تطور النماذج التوليدية الحديثة بسرعة كبيرة، وأصبحت من التقنيات الرئيسية في مجال الذكاء الاصطناعي لمهام مثل توليد النصوص والصور والأصوات، ومع ذل، ركزت معظم الأبحاث على الجوانب التقنية والحاسوبية لهذه النماذج، في حين لم يتلق الأساس الإحصائي والاحتمالي الذي تعتمد عليه هذه النماذج نفس المستوى من الاهتمام، هذا الأساس العلمي هو أساس النماذج التوليدية؛ حيث يحدد كيفية تعلم هذه النماذج وتوليد مخرجات جديدة وتقييم جودتها (Bishop & Bishop, 2024؛ Murphy, 2023؛Goodfellow et al., 2016).

هناك حاجة ماسة إلى فهم أعمق للأسس الإحصائية والاحتمالية التي تدور حول النماذج التوليدية الحديثة، هذا الفهم سيساعد في توضيح كيفية استخدام مفاهيم الاستدلال الإحصائي والنمذجة الاحتمالية في بناء هذه النماذج؛ بالإضافة إلى ذلك من المهم تحليل وتقويم أداء هذه النماذج باستخدام المقاييس الإحصائية الحديثة؛ بهذه الطريقة يمكننا تفسير القدرات والقيود الخاصة بهذه النماذج بشكل أفضل وتحديد التحديات التي قد تواجهها عند استخدامها في مختلف التطبيقات (Murphy, 2023؛ Bishop & Bishop, 2024؛Borji, 2019).

وفي البيئة السعودية، تبرز أهمية هذه المشكلة في ظل التوسع المتسارع الذي تشهده الشركات في تبني تقنيات الذكاء الاصطناعي دعمًا للتحول الرقمي وتحسين كفاءة الأداء واتخاذ القرار؛ بما يتوافق مع مستهدفات رؤية المملكة العربية السعودية 2030، ورغم تزايد الاهتمام بتطبيق النماذج التوليدية الحديثة في مجالات تحليل البيانات، وإعداد التقارير، والتنبؤ، وأتمتة العمليات؛ فإن توظيف هذه النماذج داخل الشركات لا يزال يواجه تحديات تتعلق بمدى اعتمادها على أسس إحصائية واحتمالية سليمة تضمن دقة النتائج وموثوقيتها وقابليتها للتفسير.

كما أن الأدبيات العربية، ولا سيما الدراسات التي تناولت الشركات السعودية، ما زالت محدودة في تحليل العلاقة بين الأسس الإحصائية الاحتمالية وبناء النماذج التوليدية وانعكاس ذلك على تقييم أداء الشركات، الأمر الذي يبرز الحاجة إلى مراجعة علمية متعمقة للأدبيات المتخصصة؛ لتحديد أبرز الاتجاهات البحثية، والكشف عن الفجوات المعرفية، وبيان الدور الذي تؤديه الأسس الإحصائية في تعزيز كفاءة النماذج التوليدية ودقتها عند استخدامها في تقييم أداء الشركات السعودية.

التساؤل الرئيس:

كيف تسهم الأسس الاحتمالية والإحصائية في بناء النماذج التوليدية الحديثة (Generative AI) وتقييم أدائها، وما أبرز التحديات المرتبطة بذلك؟

الأهمية العلمية

أولًا: الأهمية النظرية

تتمثل الأهمية النظرية لهذا الموضوع في أنه يسلط الضوء على الأسس الاحتمالية والإحصائية للنماذج التوليدية الحديثة، ويبرز دور الإحصاء المحوري في تصميمها وتطويرها، يساهم هذا في

تعميق الفهم العلمي للعلاقة بين الإحصاء والذكاء الاصطناعي، وهذا الموضوع يوضح المفاهيم الإحصائية التي تعتمد عليها النماذج التوليدية، مثل النمذجة الاحتمالية والاستدلال الإحصائي والتقدير الإحصائي والمتغيرات الكامنة، وهذا يثري الأدبيات العلمية في مجال الإحصاء وعلوم البيانات والذكاء الاصطناعي، كما يستعرض أحدث الاتجاهات البحثية في هذا المجال؛ مما يوفر أساسًا معرفيًا للباحثين للاستفادة منه في الدراسات المستقبلية.

ثانيًا: الأهمية التطبيقية

تكمن الأهمية التطبيقية لهذا الموضوع في توضيح كيفية استخدام الأساليب الإحصائية في بناء النماذج التوليدية وتقييم أدائها بدقة، ويساعد هذا في تحسين جودة المخرجات وزيادة موثوقيتها، كما يبرز دور الإحصاء في معالجة التحديات المرتبطة بهذه النماذج، مثل تقدير عدم اليقين والكشف عن التحيزات وقياس قدرة النماذج على التعميم وتقييم جودة البيانات المولدة، ومن المتوقع أن تسهم نتائج هذا العمل في تحسين استخدام النماذج التوليدية في مجالات متعددة مثل الرعاية الصحية والقطاع المالي والتعليم والأمن السيبراني وتحليل.

الأهداف البحثية

هذا البحث يسعى لدراسة النماذج التوليدية الحديثة من زاوية إحصائية، سوف يُحلل الأسس الاحتمالية التي تعتمد عليها هذه النماذج، ويستعرض المنهجيات الإحصائية التي تُستخدم في بنائها (Bishop & Bishop, 2024؛Murphy, 2023)؛ بالإضافة إلى ذلك سوف يُقيّم أداء هذه النماذج وفقًا للمقاييس الإحصائية الحديثة (Goodfellow et al., 2016؛Borji, 2019)، وسوف يتناول البحث أيضًا التحديات الرئيسية والاتجاهات البحثية المستقبلية في هذا المجال؛ لتقديم رؤية شاملة حول تطور النماذج التوليدية الحديثة (Bommasani et al., 2021؛ OpenAI, 2023).

الأهداف الفرعية

1. تحديد دور النماذج التوليدية الحديثة القائمة على مبادئ الإحصاء الاحتمالي في تقييم الأداء.
2. الكشف عن أسس الإحصاء الاحتمالي المُستخدمة في بناء النماذج التوليدية الحديثة.
3. توضيح النماذج التوليدية الحديثة المُستخدمة في تقييم أداء الشركات.
4. الكشف عن التحديات التي تواجه تطبيق النماذج التوليدية الحديثة المُستخدمة في تقييم أداء الشركات.
5. طرح أبرز تجارب الشركات العالمية والدولية في تطبيق استخدام النماذج التوليدية الحديثة لتقييم أداء الشركات.
6. تقديم تصور مُقترح لتطوير استخدام النماذج التوليدية في تقييم أداء الشركات السعودية.

أدوات الدراسة

هذه الدراسة تعتمد على مجموعة من الأدوات العلمية التي تتناسب مع طبيعتها النظرية، وتشمل هذه الأدوات ما يلي:

1. نعتمد على المراجع العلمية المتخصصة، مثل الكتب الأكاديمية في الإحصاء والذكاء الاصطناعي، ولا سيما المراجع المتعلقة بالنمذجة الاحتمالية، والاستدلال الإحصائي، والنماذج التوليدية الحديثة.

2. نستفيد من الدوريات العلمية المحكمة، من خلال مراجعة وتحليل أحدث الأبحاث المنشورة في المجلات العلمية العالمية ذات الصلة بمجالات الإحصاء، والتعلم الآلي، والذكاء الاصطناعي، بهدف الوقوف على أحدث التطورات والاتجاهات البحثية.
3. نحضر وقائع المؤتمرات العلمية الدولية، للاستفادة من الدراسات الحديثة التي تناولت النماذج التوليدية وأساليب تقييمها من المنظور الإحصائي.
4. نعتمد على قواعد البيانات العلمية الإلكترونية، مثل (IEEE Xplore، وSpringerLink، وScienceDirect، وWiley Online Library، وACM Digital Library، وGoogle Scholar)، وذلك للوصول إلى أحدث الدراسات العلمية ذات الصلة بموضوع البحث.
5. نتبع أسلوب التحليل والمقارنة، وذلك بتحليل الأدبيات العلمية ومقارنة النماذج التوليدية الحديثة؛ من حيث الأسس الاحتمالية، والمنهجيات الإحصائية، وأساليب تقييم الأداء، واستخلاص النتائج والتوجهات البحثية المستقبلية.

منهجية الدراسة

1. اعتمدت هذه الدراسة على المنهج الوصفي التحليلي بوصفه المنهج الأنسب؛ لتحقيق أهدافها، وذلك من خلال وصف وتحليل الأدبيات العلمية المتعلقة بالنماذج التوليدية الحديثة (Generative AI) من منظور إحصائي، مع التركيز على الأسس الاحتمالية التي تقوم عليها وآليات تقييم أدائها (Creswell & Creswell, 2023؛ Murphy, 2023).
2. كما استندت الدراسة إلى منهج المراجعة العلمية المنهجية للأدبيات (Literature Review)؛ من خلال جمع وتحليل الدراسات والأبحاث العلمية الحديثة المنشورة في المجلات المحكمة والمؤتمرات العلمية وقواعد البيانات الأكاديمية؛ بهدف استعراض

التطور العلمي للنماذج التوليدية، وتحليل الأساليب الإحصائية المستخدمة في بنائها، وتقييم المقاييس المعتمدة للحكم على أدائها (Snyder, 2019؛ Kitchenham & Charters, 2007).

3. واعتمدت الدراسة أيضًا على المنهج المقارن عند تحليل ومقارنة أبرز النماذج التوليدية الحديثة، مثل الشبكات التنافسية التوليدية(GANs) ، والمشفرات التباينية(VAEs) ، ونماذج الانتشار(Diffusion Models) ، والنماذج اللغوية الكبيرة(LLMs) ، وذلك من حيث الأسس الاحتمالية، وآليات التعلم، وأساليب التقدير الإحصائي، ومقاييس تقييم الأداء، وأبرز المزايا والتحديات المرتبطة بكل نموذج (Goodfellow et al., 2014؛ Kingma & Welling, 2014؛ Ho et al., 2020 ؛ Brown et al., 2020).

4. وفي مرحلة التحليل، جرى استخلاص النتائج من خلال الربط بين المفاهيم الإحصائية والنظريات الاحتمالية والتطبيقات الحديثة للنماذج التوليدية، مع تقديم تحليل نقدي لأهم التحديات الإحصائية والاتجاهات البحثية المستقبلية؛ بما يسهم في تقديم رؤية علمية متكاملة حول دور الإحصاء في تطوير هذه النماذج وتحسين كفاءتها (Bishop & Bishop, 2024؛ Murphy, 2023؛ Bommasani et al., 2021).

الإطار النظري للدراسة

نظرية الاحتمالات هي الأساس الرياضي الذي تستند إليه النماذج التوليدية؛ حيث توفر الإطار العلمي لوصف الظواهر العشوائية وتمثيل عدم اليقين في البيانات، النماذج التوليدية تعتمد على افتراض أن البيانات المشاهدة تمثل عينات مأخوذة من توزيع احتمالي غير معروف، وتسعى إلى تقدير هذا التوزيع بأكبر قدر ممكن من الدقة (Murphy, 2023؛ Bishop & Bishop, 2024).

النمذجة الاحتمالية هي عملية بناء نموذج رياضي يصف كيفية توليد البيانات اعتمادًا على توزيعات احتمالية محددة، هذا النوع من النمذجة يختلف عن النماذج الحتمية؛ حيث يأخذ في الاعتبار الطبيعة العشوائية للبيانات وعدم اليقين المصاحب لها (Murphy, 2023؛ Bishop & Bishop, 2024).

النموذج التوليدي يبني نموذجًا احتماليًا يحاول محاكاة التوزيع الحقيقي للبيانات؛ بحيث يمكن استخدامه في توليد بيانات جديدة أو استنتاج خصائص البيانات غير المشاهدة، النمذجة الاحتمالية تمتاز بعدة مزايا، منها تمثيل عدم اليقين بصورة رياضية، والتعامل مع البيانات الناقصة أو غير الكاملة، وإمكانية تفسير النتائج إحصائيًا، والمرونة في دمج المعرفة السابقة ضمن عملية التعلم (Murphy, 2023؛Goodfellow et al., 2016).

تقدير المعلمات التي تصف التوزيع الاحتمالي للبيانات هو جزء أساسي في بناء النماذج التوليدية، ويستخدم في ذلك مجموعة من الأساليب الإحصائية، مثل أسلوب الإمكان الأعظم والتقدير البايزي، أسلوب الإمكان الأعظم يهدف إلى إيجاد قيم المعلمات التي تجعل البيانات المشاهدة أكثر احتمالًا وفق النموذج الإحصائي المفترض (Bishop & Bishop, 2024؛ Casella & Berger, 2021).

الاستدلال البايزي يمثل أحد أهم المفاهيم الإحصائية المستخدمة في النماذج التوليدية الحديثة، حيث يعتمد على تحديث الاعتقاد حول معلمات النموذج كلما توفرت بيانات جديدة، المتغيرات الكامنة هي متغيرات غير قابلة للملاحظة مباشرة، ولكنها تمثل العوامل الداخلية التي تفسر كيفية توليد البيانات (Murphy, 2023؛ Bishop & Bishop, 2024؛ Gelman et al., 2021).

الإطار المفاهيمي

يشهد مجال الذكاء الاصطناعي تطورًا كبيرًا في العقد الأخير؛ حيث انتقل الاهتمام من تصميم أنظمة قادرة على التصنيف والتنبؤ إلى تطوير نماذج تمتلك القدرة على تعلم الأنماط الكامنة في البيانات وإنتاج بيانات جديدة تحاكي البيانات الأصلية، ويستند هذا التحول إلى تطور كبير في النماذج الاحتمالية والأساليب الإحصائية، ويهدف هذا المحور إلى توضيح المفاهيم الأساسية المتعلقة بالنماذج التوليدية الحديثة، واستعراض تطورها التاريخي، وخصائصها، وبيان الفروق الجوهرية بينها وبين النماذج التمييزية (Goodfellow et al., 2016؛ Murphy, 2023 ؛ Bishop & Bishop, 2024).

1-مفهوم الذكاء الاصطناعي التوليدي

يشير الذكاء الاصطناعي التوليدي إلى مجموعة من النماذج والخوارزميات القادرة على تعلم التوزيع الاحتمالي للبيانات الأصلية، ثم استخدام هذا التعلم لتوليد بيانات جديدة تتشابه مع البيانات الحقيقية في خصائصها الإحصائية، يختلف هذا النوع من الذكاء الاصطناعي عن النظم التقليدية التي تقتصر وظيفتها على التصنيف أو التنبؤ، ومن المنظور الإحصائي يمكن النظر إلى النماذج التوليدية على أنها نماذج تهدف إلى تقدير التوزيع الاحتمالي المشترك للمتغيرات. يعتمد نجاح هذه النماذج على مدى قدرتها على تمثيل البنية الإحصائية الحقيقية للبيانات، وتوسعت تطبيقات الذكاء الاصطناعي التوليدي لتشمل إنتاج النصوص، وإنشاء الصور ومقاطع الفيديو، وتوليد الأصوات، وتصميم الأدوية، وإنتاج البيانات الاصطناعية (Goodfellow et al., 2016؛ Brown et al., 2020؛ OpenAI, 2023؛ Murphy, 2023).

مجتمع الدراسة

يتمثل مجتمع الدراسة في الأبحاث والكتب العلمية الحديثة المتعلقة بالنماذج التوليدية، مثل المقالات الأكاديمية، والدراسات المحكمة، وأوراق المؤتمرات، والتقارير البحثية، ركزت هذه الدراسات على الأسس الاحتمالية والإحصائية لهذه النماذج، وأساليب بنائها، وتقييم أدائها، وتطبيقاتها المختلفة.

تتضمن الدراسة بشكل خاص الأبحاث الحديثة المنشورة بين عامي 2020 و2026، نظرًا للتطور السريع في هذا المجال في الآونة الأخيرة، تستخدم الدراسة أيضًا بعض المراجع الأساسية القديمة التي وضعت الأسس النظرية والإحصائية للنماذج التوليدية.

تم اختيار عينة الدراسة بعناية من الأبحاث العلمية الحديثة ذات الصلة بالنماذج التوليدية من منظور إحصائي، وتم اختيار الأبحاث بناءً على عدة معايير، أهمها الارتباط المباشر بموضوع الدراسة، وحداثة النشر، وجودة المجلة أو المؤتمر الذي نُشرت فيه، وركزت الدراسة على الأبحاث التي تتناول الأسس الاحتمالية، والاستدلال الإحصائي، وتقييم الأداء، والتطبيقات الحديثة لهذه النماذج.

حدود الدراسة

أولًا: نطاق الموضوع

هذه الدراسة تركز فقط على النماذج التوليدية الحديثة المعروفة بـ الذكاء الاصطناعي التوليدي من وجهة نظر إحصائية، وستركز الدراسة على الأسس الاحتمالية التي تقوم عليها هذه النماذج، والطرق الإحصائية التي تستخدم في بنائها، وكيفية تقييم أدائها؛ بالإضافة إلى أبرز تطبيقاتها

والتحديات التي تواجهها والاتجاهات البحثية المستقبلية، ولن نخوض في تفاصيل البرمجة أو الهندسة المتعلقة بتطوير هذه الخوارزميات.

ثانيًا: الإطار الزمني

الدراسة تركز على الأبحاث الحديثة المنشورة بين عامي 2020 و2026، كما استخدمنا بعض المراجع الأساسية السابقة التي ساهمت في وضع الأسس الإحصائية والنظرية للنماذج التوليدية.

ثالثًا: النطاق المكاني

الدراسة تقتصر على الأبحاث المنشورة في المجلات والدوريات المحكمة وقواعد البيانات الأكاديمية العالمية، لم نقم بتقييد البحث بدولة أو مؤسسة بحثية معينة بسبب الطبيعة النظرية للدراسة.

رابعًا: المنهجية

استخدمنا في هذه الدراسة المنهج الوصفي التحليلي ومنهج المراجعة العلمية للأدبيات، كما استفدنا من المنهج المقارن لتحليل ومقارنة أبرز النماذج التوليدية الحديثة من المنظور الإحصائي، لم نقم بإجراء دراسة ميدانية أو تجريبية أو تطبيق عملي على بيانات حقيقية.

الدراسات السابقة

كثير من الباحثين درسوا موضوع النماذج التوليدية الحديثة من زوايا مختلفة، بعض الدراسات ركزت على الأسس النظرية لهذه النماذج، بينما ركزت دراسات أخرى على تطبيقاتها العملية أو تقييم أدائها أو تحليل التحديات المرتبطة بها، هنا عرض لأبرز الدراسات ذات الصلة:

الدراسة (Balasubramaniam ، 2024) بعنوان: الطريق إلى الأمام: الاتجاهات الناشئة والقضايا غير المحلولة وملاحظات الختام في الذكاء الاصطناعي التوليدي – مراجعة شاملة.

هدفت هذه الدراسة إلى تقديم نظرة شاملة حول تطور الذكاء الاصطناعي التوليدي، واستعراض أهم النماذج التوليدية الحديثة، مثل الشبكات التنافسية التوليدية والمشفرات التباينية ونماذج الانتشار، كما ناقشت تطبيقاتها في مجالات متعددة مثل الرؤية الحاسوبية ومعالجة اللغة الطبيعية والرعاية الصحية؛ بالإضافة إلى مناقشة التحديات الأخلاقية والبحثية في المستقبل.

أظهرت نتائج هذه الدراسة أن النماذج التوليدية أصبحت تمثل اتجاهًا هامًا في مجال الذكاء الاصطناعي؛ إلا أن هناك حاجة ماسة لتطوير أساليب أكثر موثوقية لتقييم الأداء، وتحسين قابلية التفسير، ومعالجة مشكلات التحيز وعدم اليقين.

دراسة (Megahed وآخرون ، 2024) بعنوان: مراجعة شاملة للشبكات التنافسية التوليدية: الأسس والتطبيقات والتحديات.

ركزت هذه الدراسة على الشبكات التنافسية التوليدية، موضحة الأسس النظرية التي تقوم عليها، وآليات التدريب، وأهم التطبيقات في توليد الصور والبيانات، كما ناقشت أبرز المشكلات المرتبطة بها، مثل عدم استقرار التدريب وانهيار الأنماط؛ بالإضافة إلى استعراض أهم مقاييس تقييم الأداء.

أشارت الدراسة إلى أن نجاح نماذج الشبكات التنافسية التوليدية يرتبط بتحسين الأساليب الإحصائية المستخدمة في التدريب والتقييم، مع التأكيد على الحاجة إلى تطوير مقاييس تقييم أكثر دقة وارتباطًا بالإدراك البشري.

دراسة (Chen وآخرون ، 2024) بعنوان نماذج الانتشار التوليدية: مسح للتطورات النظرية الحالية.

هدفت هذه الدراسة إلى مراجعة التطورات النظرية في نماذج الانتشار، مع التركيز على الأساس الرياضي والإحصائي لهذه النماذج، وآليات التوليد، والمقاييس المستخدمة في تقييم الأداء، كما استعرضت أبرز التطبيقات في الصور والنصوص والبيانات الحيوية.

أوضحت النتائج أن نماذج الانتشار تتميز باستقرار عملية التدريب وجودة المخرجات مقارنة بعدد من النماذج التوليدية الأخرى؛ إلا أن تطوير الأسس النظرية والمقاييس الإحصائية ما يزال يمثل مجالًا بحثيًا نشطًا.

دراسة (Cai,2024) بعنوان: تعزيز قدرات النماذج التوليدية من خلال دمج المشفرات التباينية مع الشبكات التنافسية التوليدية (مراجعة).

استعرضت هذه الدراسة محاولات دمج المشفرات التباينية مع الشبكات التنافسية التوليدية؛ بهدف الاستفادة من المزايا الإحصائية لكلا النموذجين، وتحسين جودة البيانات المولدة واستقرار عملية التدريب.

أظهرت النتائج أن النماذج الهجينة تحقق توازنًا أفضل بين جودة المخرجات والتمثيل الاحتمالي للبيانات، مقارنة باستخدام كل نموذج بصورة منفردة.

دراسة (Chen وMei وFan ، 2024) بعنوان: نظرة عامة على نماذج الانتشار: التطبيقات والانتشار الموجه والأسس الإحصائية والتأمل.

هدفت هذه الدراسة إلى تحليل الأسس النظرية والإحصائية لنماذج الانتشار، مع التركيز على معدلات التقارب الإحصائي، وأساليب التحسين، والتطبيقات الحديثة في مجالات الرؤية الحاسوبية والصوت والبيانات الحيوية.

أكدت الدراسة أن نماذج الانتشار تمثل اتجاهًا واعدًا في الذكاء الاصطناعي التوليدي؛ إلا أن فهم خصائصها الإحصائية بصورة أعمق يمثل أحد أهم التحديات البحثية المستقبلية.

من خلال المراجعة للدراسات السابقة، يتضح أنها ركزت على الجوانب النظرية والتطبيقية للنماذج التوليدية، واستعرضت تطورها وأبرز تطبيقاتها والتحديات التي تواجهها، كما تناولت بعض الدراسات نماذج محددة مثل الشبكات التنافسية التوليدية أو نماذج الانتشار بصورة متخصصة، ومع ذلك، فإن معظم هذه الدراسات جاءت من منظور الذكاء الاصطناعي وعلوم الحاسب، بينما لم تحظ الأسس الاحتمالية والإحصائية للنماذج التوليدية، وآليات تقييم الأداء باستخدام المقاييس الإحصائية الحديثة، بالقدر نفسه من التحليل المتكامل.

هنا تكمن أهمية الدراسة الحالية؛ حيث تتميز بتركيزها على تقديم رؤية إحصائية شاملة للنماذج التوليدية الحديثة؛ من خلال تحليل الأسس الاحتمالية التي تقوم عليها، واستعراض أساليب التقدير والاستدلال الإحصائي، ومناقشة المقاييس الإحصائية المستخدمة في تقييم الأداء؛ بما يسهم في سد جانب من الفجوة المعرفية في هذا المجال.

تقسيمة الدراسة المقترحة

لتحقيق أهداف الدراسة والإجابة عن تساؤلاتها، تم تقسيم الدراسة إلى مقدمة وخمسة محاور رئيسة، وهي:

المحور الأول: الإطار المفاهيمي للنماذج التوليدية الحديثة

1-مفهوم الذكاء الاصطناعي التوليدي

2-التطور التاريخي للنماذج التوليدية، وخصائصها

3-الفرق بين النماذج التوليدية والنماذج التمييزية

المحور الثاني: الأسس الاحتمالية والإحصائية للنماذج التوليدية الحديثة

1-نظرية الاحتمالات

2-النمذجة الاحتمالية

3-طرق التقدير الإحصائي

4-الاستدلال البايزي

5-المتغيرات الكامنة

6-قياس عدم اليقين

المحور الثالث: النماذج التوليدية الحديثة من المنظور الإحصائي

1-النماذج التوليدية، مثل نماذج الخلط الغاوسي والمشفرات التباينية والشبكات التنافسية التوليدية ونماذج الانتشار والنماذج اللغوية الكبيرة، مع إجراء مقارنة إحصائية بينها.

المحور الرابع: تقييم أداء النماذج التوليدية من المنظور الإحصائي

1-مفهوم تقييم الأداء

2-المقاييس الإحصائية المستخدمة في تقييم النماذج

3-تقييم جودة البيانات المولدة

4-قياس عدم اليقين

5-تحليل التحيز

6-قابلية التعميم

المحور الخامس: التطبيقات العملية والتحديات والاتجاهات المستقبلية للنماذج التوليدية الحديثة

1-أبرز التطبيقات في مختلف المجالات

2-التحديات الإحصائية والتقنية والأخلاقية

3-الاتجاهات البحثية المستقبلية

4-دور الإحصاء في تطوير النماذج التوليدية

خاتمة

في ظل التغيرات السريعة التي تشهدها مجال الذكاء الاصطناعي، ظهرت النماذج التوليدية الحديثة كأحد أكثر التقنيات تأثيرًا في مختلف المجالات العلمية والتطبيقية، وذلك لقدرتها على تعلم الأنماط المخفية في البيانات وإنتاج بيانات جديدة تشبه الواقع بدرجة عالية من الدقة، ويرتكز نجاح هذه النماذج على أسس احتمالية وإحصائية متقدمة؛ مما يجعل علم الإحصاء عنصرًا أساسيًا في تصميمها وتطويرها وتقييم أدائها.

من هذا المنظار، جاءت هذه الدراسة لتقديم رؤية علمية شاملة للنماذج التوليدية الحديثة من منظور إحصائي؛ من خلال استعراض الأسس الاحتمالية التي تقوم عليها، وتحليل أبرز النماذج التوليدية، ومناقشة المقاييس الإحصائية المستخدمة في تقييم أدائها؛ بالإضافة إلى استعراض تطبيقاتها العملية والتحديات والاتجاهات البحثية المستقبلية، كما تهدف الدراسة إلى تسليط الضوء على الدور المتنامي للإحصاء في دعم تطوير تقنيات الذكاء الاصطناعي التوليدي، وتعزيز موثوقيتها، وتحسين قدرتها على التعامل مع البيانات.

من المتوقع أن تساهم هذه الدراسة في إثراء الجانب النظري في مجال الإحصاء والذكاء الاصطناعي، وأن تمثل مرجعًا علميًا للباحثين والدارسين المهتمين بالنماذج التوليدية الحديثة، بالإضافة إلى فتح آفاق لدراسات مستقبلية تتناول تطوير الأساليب الإحصائية الخاصة بالنمذجة الاحتمالية، والاستدلال البايزي، وتقييم الأداء؛ بما يواكب التطور المتسارع في تقنيات الذكاء الاصطناعي وتطبيقاته.

قائمة المراجع

أولًا: المراجع العربية

1. أبو زيد، أحمد محمد .(2024) .**الذكاء الاصطناعي التوليدي: المفاهيم والتطبيقات والتحديات** .القاهرة: دار الفكر العربي.
2. السيد، محمد إبراهيم .(2023) .**الأساليب الإحصائية الحديثة في تحليل البيانات الضخمة** .الإسكندرية: دار المعرفة الجامعية.
3. عبد الحميد، محمود أحمد .(2024) .**الإحصاء والذكاء الاصطناعي: التكامل والتطبيقات المعاصرة** .عمان: دار المسيرة للنشر والتوزيع.

ثانيًا: المراجع الأجنبية

1. Balasubramaniam, S. (2024). **The Road Ahead: Emerging Trends, Unresolved Issues, and Concluding Remarks in Generative AI—A Comprehensive Review**. Applied and Computational Engineering, 2024, 1–31.

2. Bishop, C. M. (2006). **Pattern Recognition and Machine Learning**. Springer.
3. Chen, M., Mei, S., & Fan, J. (2024). **An Overview of Diffusion Models: Applications, Guided Generation, Statistical Rates and Optimization**. arXiv preprint arXiv:2404.07771.
4. Goodfellow, I., Bengio, Y., & Courville, A. (2016). **Deep Learning**. MIT Press.
5. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). **Generative Adversarial Networks**. Advances in Neural Information Processing Systems (NeurIPS).
6. Jordan, M. I. (2004). **Graphical Models**. Statistical Science, 19(1), 140–155.
7. Kingma, D. P., & Welling, M. (2014). **Auto-Encoding Variational Bayes**. International Conference on Learning Representations (ICLR).
8. Koller, D., & Friedman, N. (2009). **Probabilistic Graphical Models: Principles and Techniques**. MIT Press.
9. MacKay, D. J. C. (2003). **Information Theory, Inference, and Learning Algorithms**. Cambridge University Press.

10. Murphy, K. P. (2022). **Probabilistic Machine Learning: An Introduction**. MIT Press.

11. Murphy, K. P. (2023). **Probabilistic Machine Learning: Advanced Topics**. MIT Press.

12. Pearl, J. (2009). **Causality: Models, Reasoning, and Inference** (2nd ed.). Cambridge University Press.

13. Rasmussen, C. E., & Williams, C. K. I. (2006). **Gaussian Processes for Machine Learning**. MIT Press.

14. Rezende, D. J., Mohamed, S., & Wierstra, D. (2014). **Stochastic Backpropagation and Approximate Inference in Deep Generative Models**. International Conference on Machine Learning (ICML).

15. Shalev-Shwartz, S., & Ben-David, S. (2014). **Understanding Machine Learning: From Theory to Algorithms**. Cambridge University Press.

16. Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). **Attention Is All You Need**. Advances in Neural Information Processing Systems (NeurIPS).