

التكليف الثامن نقد التحليل الإحصائي

اسم الطالبة: زكية السلمي
المقرر: دراسات نقدية
المرحلة: دكتوراه فلسفة الإدارة
جامعة ميدأوشن



الدراسة محل النقد:

الذكاء الاصطناعي وتأثيره في التدقيق الداخلي: دراسة استطلاعية لآراء عينة من المدققين في جامعة الموصل

لا يمثل التحليل الإحصائي عملية حسابية تأتي في نهاية البحث، بل هو حلقة مركزية في إنتاج الدليل العلمي؛ إذ يحوّل الباحث من خلاله الخبرات والتصورات والممارسات التنظيمية إلى مؤشرات رقمية، ثم يعيد تفسيرها بوصفها معرفة يمكن الاستناد إليها في الحكم واتخاذ القرار. ومن منظور فلسفة الإدارة، لا تكمن قيمة التحليل في كثرة الاختبارات المستخدمة، وإنما في مدى انسجامها مع: طبيعة الظاهرة الإدارية.

مستوى قياس البيانات.

أسئلة الدراسة وفرضياتها.

جودة أداة القياس.

حدود التفسير والتعميم.

أرى أن الإحصاء لا يتحدث بذاته؛ فالرقم يكتسب معناه من السؤال الذي أنتجه، والأداة التي قاست الظاهرة، والنظرية التي استخدمت في تفسيره.

لا يُقاس التحليل الجيد بدقة الحساب فقط، بل بمدى انضباط الاستنتاج بحدود البيانات التي بُني عليها.

هدفت الدراسة إلى التعرف على أهمية تطبيق الذكاء الاصطناعي في تقليل الوقت اللازم لإتمام التدقيق الداخلي، وبيان تأثيره في جمع أدلة الإثبات.

واعتمد الباحثان المنهج الوصفي والتحليلي والاستقرائي، واختارا مدققي جامعة الموصل مجتمعًا للدراسة. وُزعت الاستبانة على 52 مدققًا، واستُردت 51 استبانة وحُللت باستخدام برنامج SPSS.

تنسجم الطبيعة الكمية للبيانات مع استخدام التحليل الإحصائي، لكن وصف الدراسة بأنها استطلاعية واستقرائية يثير تساؤلًا حول عمق الاستكشاف؛ لأن الاعتماد النهائي كان على استبانة مغلقة وتحليل رقمي. كما أن استجابة 51 مشاركًا من مجتمع يبلغ 52 شخصًا تجعل الدراسة أقرب إلى المسح شبه الشامل لمجتمع محدود، وليس إلى عينة احتمالية واسعة.

وتزداد الإشكالية لأن الدراسة ذكرت أن وحدات التدقيق لا تستخدم الذكاء الاصطناعي فعليًا، بل تعتمد البرامج التقليدية؛ وبذلك تصبح إجابات المشاركين أقرب إلى قياس تصوراتهم عن الأثر المتوقع، لا قياس أثر تطبيق قائم بالفعل.

أرى أن هناك فجوة بين لغة الدراسة التي تتحدث عن أثر الاستخدام والواقع الميداني الذي يشير إلى غياب الاستخدام الفعلي. وكان من الأدق استخدام تعبير مثل: الاتجاهات نحو الأثر المتوقع لتطبيق الذكاء الاصطناعي في التدقيق الداخلي. الدراسة صالحة لاستكشاف التصورات المهنية، لكنها لا تقدم وحدها دليلًا ميدانيًا كافيًا على أثر تطبيق فعلي.

اعتمدت الدراسة مقياس ليكرت الخماسي، وأعطت بدائل الاستجابة قيمًا من 1 إلى 5، ثم قسمت المتوسطات إلى ثلاثة مستويات: من 1 إلى 2.33: درجة منخفضة.

من 2.34 إلى 3.67: درجة متوسطة.

من 3.68 إلى 5: درجة مرتفعة.

تُعد استجابات ليكرت في أصلها بيانات رتبية، لأن فئاتها مرتبة، لكن تساوي المسافات النفسية بين هذه الفئات ليس أمرًا مؤكدًا. وقد استخدمت الدراسة المتوسطات والانحرافات المعيارية والارتباط والانحدار، وهو أمر شائع في البحوث الإدارية عندما تُدمج عدة فقرات في مقياس مركب.

لكن الدراسة لم توضح:

هل تحقق الاعتدال في توزيع البيانات؟

هل فُحصت القيم المتطرفة؟

هل تحققت خطية العلاقات؟

هل كانت البيانات ملائمة للاختبارات المعلمية؟

ما الأساس المستخدم في تقسيم المتوسطات إلى ثلاث فئات؟

لا أعترض على استخدام المتوسط والانحدار لمجرد أن البيانات مستمدة من ليكرت، لكنني أرى أن قبول هذه الإجراءات مشروط بتبريرها والتحقق من افتراضاتها، لا بمجرد شيوعها في الدراسات السابقة.

الاختبار الإحصائي لا يصبح مناسبًا لأنه مألوف، بل لأنه يتوافق مع مستوى القياس وافتراضات البيانات.

استخدمت الدراسة التكرارات والنسب لوصف المشاركين. وأظهرت النتائج أن 94.1% يحملون درجة البكالوريوس، و78.4% من تخصص المحاسبة، و54.9% يعملون بمسمى مدقق أقدم، كما بلغت نسبة أصحاب الخبرة من خمس إلى أقل من عشر سنوات 66.7%. كما بلغ المتوسط العام لمحور الذكاء الاصطناعي 4.421 بانحراف معياري 0.440، بما يشير إلى اتفاق مرتفع بين المشاركين.

جوانب القوة

ملاءمة التكرارات والنسب للخصائص الديموغرافية.

استخدام المتوسطات والانحرافات لتحديد اتجاه الاستجابات وتشتتها.

عرض النتائج على مستوى الفقرات، وليس الاكتفاء بالمعدل العام.

ارتفاع المتوسطات بشدة وندرة الاستجابات السلبية قد يعكسان اقتناعًا حقيقيًا، لكنهما قد ينتجان أيضًا عن:

صياغة الفقرات في اتجاه إيجابي واحد.

تحيز الرغبة الاجتماعية.

الانبهار بالتقنيات الحديثة.

تقييم فوائد افتراضية لم تُختبر بالممارسة.

تجانس مجتمع الدراسة وصغر حجمه.

كما لم تعرض الدراسة الوسيط أو المدى الربيعي، وهما مفيدان في قراءة بيانات ليكرت، ولم تناقش احتمال وجود انحياز الموافقة.

أرى أن المتوسط المرتفع لا يثبت وحده ارتفاع مستوى التطبيق؛ فقد يثبت ارتفاع مستوى القناعة أو التوقع فقط. ومن الضروري الفصل بين:

الاتفاق على أهمية التقنية، ووجود التقنية، وفاعلية تطبيقها.

الإحصاء الوصفي عرض اتجاهات المشاركين بوضوح، لكنه لم يفسر مصدر هذا الاتفاق المرتفع ولا علاقته بالواقع التطبيقي.

جمعت الدراسة فقرات الاستبانة في محاور، ثم حسبت لها متوسطات عامة واستخدمتها في اختبار الفرضيات. غير أن لم يتم عرض بصورة واضحة نتائج تفصيلية تتعلق بـ:
معامل ألفا كرونباخ.
صدق البناء.
ارتباط كل فقرة بالمحور.
التحليل العاملي.
اختبار التمييز بين المتغير المستقل والتابع.
قبل استخدام الارتباط والانحدار، يجب التأكد من أن الأداة تقيس متغيرات محددة ومتجانسة.
فإذا كانت بعض الفقرات تقيس التصور عن فوائد الذكاء الاصطناعي، وبعضها يقيس الاستعداد للتطبيق، وأخرى تقيس الأثر المتوقع، فإن دمجها في درجة واحدة قد يؤدي إلى غموض مفاهيمي.
والتحليل الإحصائي مهما بلغت دقته لا يستطيع معالجة خلل بدأ في بناء المفهوم أو صياغة الأداة.
أرى أن الفجوة الأساسية ليست في افتقار الدراسة إلى اختبارات أكثر تعقيداً، بل في عدم تقديم أدلة كافية على جودة القياس الذي قامت عليه هذه الاختبارات.
لا يمكن أن يكون الاستدلال أقوى من الأداة التي أنتجت بياناته؛ فصدق النتائج يبدأ من صدق المفهوم والقياس.

تناولت الفرضية الأولى أثر الذكاء الاصطناعي في تقليل الوقت اللازم لإتمام التدقيق الداخلي. واستخدمت الدراسة معامل الارتباط، الذي بلغت قيمته 0.288 عند مستوى دلالة 0.05، ثم رفضت الفرضية الصفرية وقبلت وجود أثر ذي دلالة إحصائية.

تدل القيمة 0.288 على ارتباط موجب، لكنه ضعيف نسبيًا.

والإشكالية ليست في ضعف القيمة فقط، وإنما في استخدام معامل الارتباط لإثبات الأثر.

فالارتباط يجيب عن سؤال: هل يتغير المتغيران معًا؟

بينما السببية تتطلب الإجابة عن سؤال: هل أحدث المتغير المستقل تغيرًا في المتغير التابع، مع استبعاد التفسيرات البديلة؟

ولإسناد ادعاء سببي أقوى، كان يلزم تصميم يسمح بترتيب زمني واضح، أو بيانات فعلية قبل التطبيق وبعده، أو ضبط للمتغيرات الأخرى، أو على الأقل استخدام نموذج انحدار مبرر.

أرى أن النتيجة تسمح بالقول بوجود علاقة موجبة محدودة بين تصورات المشاركين نحو الذكاء الاصطناعي وتصورهم حول تقليل الوقت، لكنها لا تسمح بالجزم بأن الذكاء الاصطناعي أحدث هذا الانخفاض فعليًا.

اختبرت الدراسة أثر الذكاء الاصطناعي في جمع أدلة الإثبات باستخدام الانحدار، وأظهرت النتائج:

معامل انحدار β بلغ 0.339.

قيمة F بلغت 4.447.

معامل تحديد R^2 بلغ 0.083.

قيمة t محسوبة بلغت 2.109.

مستوى الدلالة المستخدم 0.05.

استخدام الانحدار هنا أكثر ملاءمة من استخدام الارتباط وحده؛ لأنه يقدر مقدار التغير في المتغير التابع المرتبط بتغير المتغير المستقل. لكن معامل التحديد يوضح أن النموذج يفسر فقط 8.3% من التباين في جمع أدلة الإثبات، بينما تعود النسبة الأكبر إلى عوامل أخرى لم يتضمنها النموذج.

وقد تشمل هذه العوامل: خبرة المدقق. جودة أنظمة المعلومات. دعم الإدارة. البنية التقنية. التدريب. الضوابط المهنية. أمن المعلومات.

ولا تُذكر هذه العوامل بوصفها نتائج مثبتة في الدراسة، وإنما بوصفها احتمالات تفسيرية كان يمكن للنموذج البحثي اختبارها.

أرى أن الدراسة ركزت على تحقق الدلالة، لكنها لم تمنح القوة التفسيرية المحدودة للنموذج وزنها النقدي الكافي.

فالسؤال العلمي لا ينبغي أن يتوقف عند: هل توجد دلالة؟

بل يجب أن يمتد إلى: كم فسّر النموذج؟ وما الذي لم يفسره؟ وهل النتيجة مهمة إداريًا؟

خلصت الدراسة إلى وجود أثر للذكاء الاصطناعي في تقليل الوقت وفي جمع أدلة الإثبات، وأضافت نتائج تتعلق باكتشاف الأخطاء وتحسين جودة التدقيق.

ينبغي التمييز بين ثلاثة مستويات:

الدالة الإحصائية:

هل يصعب تفسير النتيجة بالعشوائية وفق مستوى الدالة المختار؟

حجم الأثر:

ما مقدار العلاقة أو التغير الذي يفسره المتغير؟

الدالة الإدارية:

هل هذا المقدار كافٍ لإحداث فرق مهم في واقع العمل؟

وقد أثبتت الدراسة بعض العلاقات الدالة، لكنها لم تقدم تحليلًا كافيًا للقيمة العملية، مثل:

مقدار الوقت الذي يمكن توفيره فعليًا. مقدار التحسن في دقة التدقيق. تكلفة تطبيق التقنية مقابل منافعها. أثرها في الحكم المهني للمدقق.

مخاطر الخصوصية والأمن والاعتماد المفرط.

كما أن بعض النتائج النهائية أوسع من الفرضيات التي اختُبرت مباشرة؛ إذ انتقلت الدراسة إلى الحديث عن الجودة واكتشاف الأخطاء من

غير عرض نموذج استدلال مستقل لكل نتيجة.

أرى أن الأرقام قدمت مؤشرًا أوليًا على قبول التقنية وتوقع فوائدها، لكنها لم تقدم بعد دليلًا كافيًا لتقرير فاعليتها المؤسسية الشاملة.

المحور	ما قدمته الدراسة	القصور النقدي	موقف الباحثة
طبيعة البيانات	استبانة كمية وليكرت خماسي	لم تناقش مستوى القياس وافتراضات الاختبارات	استخدام الأساليب المعلمية يحتاج إلى تبرير وفحص
الإحصاء الوصفي	تكرارات ونسب ومتوسطات وانحرافات	المتوسطات المرتفعة لم تُفسر في ضوء التحيز والسياق	النتائج تصف القناعة أكثر مما تثبت التطبيق
جودة القياس	جمع الفقرات في محاور كلية	غياب عرض واضح للثبات وصدق البناء	جودة القياس تسبق قوة الاستدلال
الفرضية الأولى	ارتباط موجب بلغ 0.288	استخدام الارتباط لإثبات الأثر	النتيجة تدعم علاقة ضعيفة لا سببية مؤكدة
الفرضية الثانية	$R^2=0.083$ و $\beta=0.339$	دلالة إحصائية مع تفسير محدود	ينبغي إبراز حجم التفسير والعوامل الغائبة
تفسير النتائج	ربط الذكاء الاصطناعي بالوقت والأدلة والجودة	بعض النتائج أوسع من النماذج المختبرة	يجب أن تظل الاستنتاجات داخل حدود الأدلة
القيمة الإدارية	توصيات بالتدريب والتطبيق	لم تُدرس التكلفة والمخاطر والجاهزية	قرار التبني يحتاج إلى أدلة تطبيقية متعددة



Thank You