

دليل وارشادات تصنيف النكات العربية

إرشادات تصنيف النكات العربية ضمن إطار بحثي في النمذجة اللغوية

يقدم هذا الدليل مجموعة من الإرشادات والمعايير المنهجية المعتمدة لتصنيف النكات العربية ضمن إطار بحثي في مجال النمذجة اللغوية ومعالجة اللغة الطبيعية. يهدف الدليل إلى توحيد فهم المعلقين لمفاهيم التصنيف، وضمان الاتساق والدقة في توصيف أنواع الدعاية العربية المختلفة، بما يعكس الاستخدام الفعلي للغة في سياقاتها الثقافية والاجتماعية المتنوعة.

تم إعداد هذا الدليل ليكون مرجعًا عمليًا ومنهجيًا يُستخدم أثناء عمليات التعليق اللغوي (Annotation)، ويساعد في إنتاج بيانات بحثية عالية الجودة قابلة للاستخدام في الدراسات الأكاديمية وتدريب النماذج اللغوية. ويركز الدليل على الوصف الموضوعي للمحتوى دون إصدار أحكام قيمة، مع مراعاة الخصوصية الثقافية للغة العربية واختلاف لهجاتها.

**NATURAL LANGUAGE
PROCESSING (NLP)**

By Osamah Hussein Al-Aini

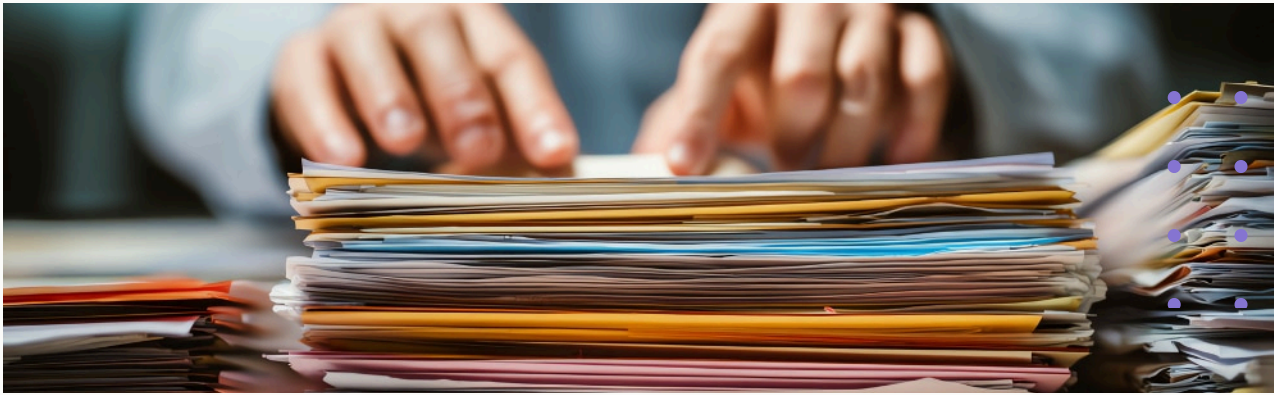
جدول المحتويات

معايير التصنيف وإطار القرار	04	مقدمة الدليل	01
آلية اتخاذ القرار في التصنيف	05	الخلفية البحثية والمنهجية	02
الأمثلة المشروحة والحالات الإشكالية	06	التعريفات والمصطلحات الأساسية	03
الاعتبارات اللغوية والثقافية وإجراءات التعليق	08	فئات تصنيف النكات العربية	04

مقدمة الدليل

الغرض من الدليل

يهدف هذا الدليل إلى توفير إطار عمل موحد ودقيق لتصنيف النكات العربية إلى ثلاث فئات رئيسية لبناء نموذج تعلم آلي قادر على التمييز بين أنماط الدعابة المختلفة في اللغة العربية. سيتم استخدام جميع الفئات في تدريب النموذج وتقييمه، حيث تساهم كل فئة في تطوير فهم شامل للفكاهة العربية بتنوعها وثنائها.



نطاق الاستخدام

بناء نماذج التعلم الآلي (LLM) لتصنيف الدعابة العربية تلقائياً
دراسة أنماط الفكاهة عبر اللهجات العربية المختلفة
تحسين قدرة الأنظمة الذكية على فهم السياقات الفكاهية
دراسة الخصائص اللغوية والثقافية للدعابة العربية

التطبيق الأساسي
البحث الأكاديمي
تطوير الأنظمة
التحليل اللغوي

الجمهور المستهدف

المعلّمون اللغويون (Annotators) الناطقون بالعربية كلغة أم
الباحثون في مجال معالجة اللغات الطبيعية
طلاب الدراسات العليا في اللسانيات الحاسوبية
متطلبات: إجادة العربية، فهم اللهجات المختلفة، إلمام بالسياق الثقافي العربي

خلفية بحثية ومنهجية

تصنيف الدعابة في الدراسات اللغوية

تُعدّ الدعابة من أكثر الظواهر اللغوية تعقيداً نظراً لاعتمادها على السياق الثقافي والمعرفة المشتركة والتلاعب اللغوي. أظهرت الدراسات أن تصنيف الفكاهة يواجه تحديات كبيرة بسبب الطبيعة الذاتية للفكاهة واختلاف إدراكها بين الأفراد. الأبحاث الحديثة في معالجة اللغات الطبيعية أثبتت أن نماذج التعلم العميق يمكنها تعلم التمييز بين أنماط الفكاهة المختلفة عند توفر بيانات مُصنفة بدقة، حيث حققت نماذج مثل CAMELBERT-DA دقة 72.11% في كشف الفكاهة العربية.

الدعابة العربية في سياق النمذجة اللغوية

الفكاهة العربية تتميز بخصائص فريدة تشمل:

- التنوع اللهجي: تختلف أساليب الدعابة باختلاف اللهجات (مصرية، خليجية، شامية، مغاربية)
- السياق الثقافي: الدعابة العربية مرتبطة بشكل عميق بالقيم والأعراف الاجتماعية
- التلاعب اللغوي: استخدام التورية والجناس والمعاني المزدوجة بشكل مكثف
- المراجع الثقافية: الاستناد إلى أحداث ومواقف ومعارف مشترك

الدعابة العربية في سياق النمذجة اللغوية

تم اختيار التصنيف الثلاثي (نظيفة/سوداء/جنسية) بناءً على:

- القابلية للتطبيق: هذا التصنيف مُستخدم في دراسات عالمية مثل CleanComedy و Naughtyformer
- الوضوح: الفئات الثلاث تقدم حدوداً واضحة نسبياً للمُعلّقين مما يسهل التصنيف
- الأهمية البحثية: تتيح دراسة أنماط لغوية وثقافية متنوعة
- التطبيقات العملية: تمكّن من بناء نماذج متخصصة لكل نوع من الدعابة

تعريفات ومصطلحات أساسية

مفهوم النكتة في اللغة العربية

النكتة هي قول أو موقف قصير يهدف إلى إثارة الضحك أو التسلية من خلال:

- المفاجأة أو القلب في التوقعات
- التلاعب اللغوي والتورية
- المبالغة أو التناقض
- السخرية أو المحاكاة الساخرة

السياق الثقافي: في الثقافة العربية، تُعتبر الدعابة وسيلة تواصل اجتماعي وأداة للتعبير عن الآراء بشكل غير مباشر.

أنواع الدعابة المستخدمة في هذا الدليل

تشمل النكات في هذه المجموعة:

- النكات القصيرة (One-liners): جملة أو جملتين
- نكات الحوار: تبادل قصير بين شخصين أو أكثر
- نكات الموقف: وصف موقف كوميدي
- النكات اللفظية: تعتمد على التورية أو اللعب بالألفاظ

توضيح المصطلحات (Clean / Dark / Dirty)

النكات النظيفة (Clean Jokes):

نكات مناسبة لجميع الأعمار والسياقات، خالية من المحتوى الجنسي أو العنف الشديد أو الإساءات المباشرة.

النكات السوداء (Dark Jokes):

نكات تتناول مواضيع حساسة أو محزنة أو مأساوية بطريقة فكاهية، مثل الموت والمرض والكوارث، لكن دون محتوى جنسي صريح.

النكات ذات المحتوى الجنسي (Dirty Jokes):

نكات تحتوي على إشارات أو تلميحات أو محتوى جنسي بشكل صريح أو ضمني.

فئات تصنيف النكات العربية في هذا البحث

1. النكات النظيفة (CLEAN)

التعريف الشامل:

نكات تعتمد على الفكاهة اللغوية أو المواقف الاجتماعية البريئة أو الملاحظات الحياتية دون تضمين محتوى جنسي أو مواضيع مرضية أو إساءات مباشرة.

ما لا تتضمنه:

- ❌ أي إشارات أو تلميحات جنسية
- ❌ الموت أو المرض أو الكوارث
- ❌ العنف أو الأذى الجسدي
- ❌ السخرية من الدين أو المعتقدات
- ❌ الإساءات المباشرة للمجموعات (عرقية، جنسية، إثنية)

الخصائص المميزة:

- ✅ التورية واللعب بالألفاظ
- ✅ المواقف الاجتماعية الكوميديّة
- ✅ المبالغات الكوميديّة البريئة
- ✅ السخرية اللطيفة من المواقف اليومية
- ✅ الملاحظات الذكية عن الحياة
- ✅ النكات المناسبة للسياقات العائلية

أمثلة على المواضيع المقبولة:

- التكنولوجيا والحياة الحديثة
- العلاقات الاجتماعية البريئة
- الحيوانات والطبيعة
- المواقف المحرّجة البسيطة
- الفروقات الثقافية بشكل محترم



فئات تصنيف النكات العربية في هذا البحث

2. النكات السوداء (DARK)

التعريف الشامل:

نكات تتعامل مع مواضيع قاتمة أو محزنة أو صادمة بأسلوب فكاهي (فكاهة المشنقة - Gallows Humor)، حيث يتم تحويل المواقف المأساوية إلى مادة كوميدية.

ما لا تتضمنه:

النكات السوداء قد تكون مزعجة أو غير مريحة، لكنها لا تتضمن محتوى جنسي. إذا احتوت على عناصر سوداء وجنسية معاً، تُصنف كجنسية

الخصائص المميزة:

- الموت والوفاة والجنازات
- الأمراض والإعاقات والحالات الطبية الخطيرة
- الحوادث والكوارث والمآسي
- المواضيع الوجودية والعدمية
- العنف (غير الجنسي)
- المواضيع المحرمة اجتماعياً (غير الجنسية)
- السخرية القاسية أو الجريئة



أمثلة على المواضيع المقبولة:

- الموت والآخرة بشكل ساخر
- الأمراض المزمنة أو الخطيرة
- الحوادث والكوارث
- الانتحار أو الأفكار الانتحارية
- فقدان الأحباء بطريقة كوميدية
- المواقف المأساوية المحولة لفكاهة

فئات تصنيف النكات العربية في هذا البحث

3. النكات ذات المحتوى الجنسي (DIRTY)

التعريف الشامل:

نكات تحتوي على محتوى جنسي صريح أو ضمني، أو تلميحات جنسية، أو إشارات للعلاقات الحميمة، أو وصف لأعضاء الجسم الجنسية.

اعتبار ثقافي مهم:

نظراً للطبيعة المحافظة للمجتمعات العربية، حتى التلميحات الجنسية الخفيفة يجب تصنيفها كجنسية. المعيار هنا أكثر صرامة من المعايير الغربية.

الخصائص المميزة:

- الأفعال الجنسية أو الممارسات الجنسية
- أعضاء الجسم الجنسية أو وظائفها
- التلميحات الجنسية أو المعاني المزدوجة الجنسية
- العلاقات الحميمة الزوجية (الجنسية)
- المواضيع الإنجابية بشكل فاضح
- الرغبة الجنسية أو الإثارة الجنسية
- المحتوى الإباحي أو شبه الإباحي

أمثلة على المواضيع المقبولة:

- الإشارات للجماع أو الممارسات الجنسية
- التلميحات للعلاقات الجنسية خارج الزواج
- وصف الأعضاء الجنسية بشكل مباشر أو غير مباشر
- النكات عن العذرية أو الخيانة الجنسية
- المواقف الجنسية المحرجة
- التورية الجنسية أو النكات الفاحشة



معايير التصنيف وإطار القرار

المبادئ العامة للتصنيف

01 المبدأ: التصنيف بناءً على المحتوى الفعلي

صنّف النكته بناءً على ما تحويه بالفعل، وليس على:

- ما إذا كانت مضحكة أم لا
- رأيك الشخصي فيها
- مدى إعجابك بها

02 المبدأ: مراعاة السياق الثقافي العربي

خذ بعين الاعتبار:

- المعايير الاجتماعية العربية
- الحساسية الدينية والثقافية
- اللهجات والفروقات الإقليمية

03 المبدأ: التركيز على المعنى المقصود

خذ بعين الاعتبار:

- المعايير الاجتماعية العربية
- الحساسية الدينية والثقافية
- اللهجات والفروقات الإقليمية

04 المبدأ: الاتساق في التصنيف

حافظ على معايير موحدة عبر جميع النكات التي تُصنّفها.

معايير التصنيف وإطار القرار

أولويات اتخاذ القرار

عند وجود أكثر من عنصر في النكتة، اتبع الأولويات التالية:

01

الأولوية الأولى: المحتوى الجنسي

إذا احتوت النكتة على أي محتوى جنسي (صريح أو ضمني) → جنسية حتى لو كانت تتضمن أيضاً موضوعات سوداء

02

الأولوية الثانية: المواضيع المرضية/المأساوية

إذا تعاملت مع الموت/المرض/الكوارث دون محتوى جنسي → سوداء

03

الأولوية الثالثة: غياب ما سبق

إذا خلت من الجنس والمواضيع المرضية → نظيفة

مثال توضيحي:

- نكتة عن علاقة زوجية تنتهي بوفاة أحد الطرفين:
- إذا كانت تتضمن إشارات جنسية → جنسية
- إذا ركزت على الموت دون جنس → سوداء
- إذا كانت عن الزواج بشكل بريء → نظيفة

معايير التصنيف وإطار القرار

حالات التداخل بين الفئات

- | | |
|--|-----------|
| التداخل 1: نظيفة ↔ سوداء | 01 |
| <ul style="list-style-type: none">• نكتة عن التقدم بالعمر → نظيفة (إذا كانت ملاحظات خفيفة)• نكتة عن الشيخوخة والموت → سوداء (إذا تناولت الفناء) | |
| التداخل 2: سوداء ↔ جنسية | 02 |
| <ul style="list-style-type: none">• نكتة عن موت مأساوي → سوداء• نكتة عن موت مع إشارات جنسية → جنسية (الأولوية للجنس) | |
| التداخل 3: نظيفة ↔ جنسية | 03 |
| <ul style="list-style-type: none">• نكتة عن الرومانسية والحب → نظيفة• نكتة عن العلاقة الحميمة → جنسية | |

قاعدة ذهبية:

- عند الشك، اختر الفئة الأكثر تقييداً (جنسية < سوداء < نظيفة)

معايير التصنيف وإطار القرار

آلية اتخاذ القرار في التصنيف (خطوات تصنيف النكته)

الخطوة 1: القراءة المتأنية (30-60 ثانية)

- اقرأ النكته كاملة بتمعن
- اقرأها مرتين إذا لزم الأمر
- تأكد من فهم اللهجة المستخدمة
- حدد نقطة الطرفة (Punchline)

الخطوة 2: الفحص الأولي (الفلاتر السريعة)

اسأل نفسك هذه الأسئلة بالترتيب:

أ. فلتر المحتوى الجنسي:

- هل تحتوي على إشارات جنسية؟
- هل هناك تلميحات أو معانٍ مزدوجة جنسية؟
- هل تتحدث عن العلاقات الحميمية؟
- إذا نعم → جنسية

ب. فلتر المحتوى المرضي/المأساوي:

- هل تتناول الموت أو الوفاة؟
- هل تتحدث عن أمراض خطيرة؟
- هل تصف حوادث أو مآسٍ؟
- إذا نعم (ولا جنس) → سوداء

ج. الفلتر النهائي:

- هل هي خالية من الجنس والمواضيع المرضية؟
- هل هي مناسبة للعائلة؟
- إذا نعم → نظيفة



معايير التصنيف وإطار القرار

آلية اتخاذ القرار في التصنيف (خطوات تصنيف النكته)

الخطوة 4: اتخاذ القرار النهائي

- اختر فئة واحدة فقط
- إذا بقي لديك شك → ضع علامة للمراجعة
- سجّل ملاحظتك إن وجدت

الخطوة 5: المراجعة الذاتية

- أعد قراءة النكته مع التصنيف المختار
- هل التصنيف منطقي؟
- هل اتبعت الأولويات الصحيحة؟
- غيّر إذا لزم الأمر قبل الإرسال

تمهيد قبل شجرة اتخاذ القرار

في الأقسام السابقة، تم توضيح فئات تصنيف النكات العربية والمعايير العامة المستخدمة للتمييز بينها. وبما أن بعض النكات قد تحتوي على عناصر متداخلة أو إشارات غير مباشرة، فقد يكون من الصعب أحيانًا اتخاذ قرار فوري اعتمادًا على التعريفات فقط.

لذلك، يقدّم هذا الدليل إطارًا إجرائيًا مبسّطًا يساعد المعلّق على اتخاذ قرار تصنيفي واضح ومتسق، من خلال سلسلة من الأسئلة المتتابة التي تعتمد على محتوى النكته ومعناها المقصود وسياقها اللغوي والثقافي. يهدف هذا الإطار إلى دعم المعلّق أثناء عملية التصنيف، وتقليل التفاوت في الأحكام بين المعلّقين المختلفين.

في الصفحة التالية، سيتم عرض شجرة اتخاذ القرار (Decision Tree)، والتي توضح بشكل بصري وتسلسلي الخطوات التي يجب اتباعها لتحديد الفئة الأنسب لكل نكته، بدءًا من قراءة النكته وفهم معناها، وصولًا إلى اختيار فئة التصنيف المناسبة.

ابدأ: اقرأ النكتة بالكامل وتأكد من فهم المعنى والسياق

هل النكتة من نوع: "نكات الكبار" او "نكات المزوجين"؟
(محتوى جنسي او تلميحات)

"لا"

"نعم"

هل النكتة تتناول مواضيع مثل الموت او الوفاة؟
امراض خطيرة؟
حوادث او كوارث؟
مواضيع مأساوية؟

جنسيه (نكته فيها تلميح جنسي)

"لا"

"نعم"

هل النكتة عن: الغباء؟
البخل؟
المحششين؟
الحماه/العم/الخال وغيره
المواقف اليوميه؟
التعليم؟
العمل؟
الضعايده؟
الحضارم؟
الدرازيين؟

النكته سودا (فكاهات سوداء)

هل فيها سخريه قاسية او عنصريه او تمييز؟
استهداف عرقي؟
سخريه من الدين؟
تنمر قاسيس؟
تمييز عرقي قاسيس؟

فكاهة نظيفة
مناسبة للجميع

سودا
سخريه قاسية/جريئة/
(عنصريه)

ملاحظات مهمة حول التصنيف

تشمل النكات ذات المحتوى الجنسي (نكات الكبار أو المتزوجين) أي نكتة تحتوي على:

- إشارات إلى العلاقات الحميمة
- معانٍ مزدوجة ذات دلالة جنسية
- تلميحات جنسية، حتى وإن كانت خفيفة أو غير مباشرة

تُصنّف النكات السوداء إلى نوعين رئيسيين:

1. نكات تتناول مواضيع مأساوية مثل الموت، المرض، الحوادث، أو الكوارث
2. نكات تعتمد على السخرية القاسية، أو التنمّر، أو العنصرية، أو التمييز الواضح

أما النكات النظيفّة فهي كل نكتة لا تندرج ضمن الفئتين السابقتين، وتكون مناسبة لجميع الفئات العمرية ولا تحتوي على محتوى حساس.

أولويات التصنيف عند تداخل الفئات

عند وجود أكثر من عنصر في النكتة الواحدة، يتم اعتماد الأولويات التالية:

1. إذا احتوت النكتة على أي محتوى جنسي، تُصنّف مباشرة كنكتة جنسية (Dirty)، حتى لو احتوت في الوقت نفسه على عناصر مأساوية أو عنصرية.

2. إذا لم يكن هناك محتوى جنسي، واحتوت النكتة على مواضيع مثل الموت أو المرض أو السخرية القاسية أو العنصرية، تُصنّف كنكتة سوداء (Dark).

3. إذا خلت النكتة من المحتوى الجنسي ومن المواضيع المأساوية أو السخرية القاسية، تُصنّف كنكتة نظيفة (Clean).

إرشادات إضافية في حالات الشك

في حال عدم وضوح التصنيف، يمكن الاستعانة بالأسئلة التالية:

- هل يمكن قول النكتة في مجلس عائلي دون حرج؟ → غالبًا نظيفة
- هل قد تُسبب إحراجًا إذا قيلت أمام الأطفال؟ → غالبًا جنسية
- هل تتناول الموت أو الفقد بأسلوب ساخر؟ → غالبًا سوداء
- هل تسخر من فئة أو مجموعة بشكل قاسٍ أو مهين؟ → غالبًا سوداء

⚠ عند التردد بين فئتين، اختر الفئة الأكثر تقييدًا، وقم بوضع علامة [مراجعة] مع كتابة سبب الاختيار لتسهيل المراجعة لاحقًا.

الأمثلة المشروحة والحالات الإشكالية

أمثلة نكات نظيفة

"بخل"

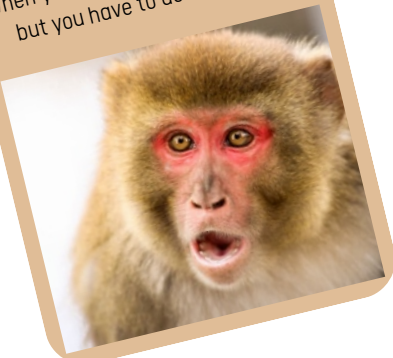
اللهجة: أردنية

النكتة: "بخيل عمل مرقة لحمة الضيوف قالوله: بس مافي لحم رد:
يعني نشاي الوزة فيه وزه."

التحليل

- النوع: نكتة البخل (موضوع شائع في الثقافة العربية)
- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ لا ❌
- سخرية قاسية أو عنصرية؟ لا ❌ (سخرية اجتماعية خفيفة)
- التصنيف: نظيفة ✅
- السبب: فكاهة اجتماعية بريئة تعتمد على التورية والمبالغة، مناسبة لجميع الأعمار

When you already know the secret,
but you have to act surprised.



”نكتة المحششين“

اللهجة: كويتية

النكتة: "في محشش اذانيه محروقه! سألوه ليش إذتك محروقه؟, قال المحشش: كنت أكوي الغترة ورن الجوال ورديت بالكواية, قالوا له والثانيه ليش محروقه؟, قال: النذل دق مرة ثانية."

التحليل

- النوع: نكتة المحششين (الغباء الكوميدي)
- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ لا ❌ (الحرق هنا خفيف وكوميدي, ليس خطيراً)
- سخرية قاسية؟ لا ❌
- التصنيف: نظيفة ✅
- السبب: موقف كوميدي يعتمد على السذاجة, الحرق جزء من الطرفة وليس أساسياً

”نكتة التعليم“

اللهجة: مصرية

مدرس يقول لطالب مشاكس مش عيب تاخذ صفر في الامتحان, فرد الطالب طالما هو عيب بتديهولي ليه!

التحليل

- النوع: نكتة المدرسة والغباء البريء
- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ لا ❌
- سخرية قاسية؟ لا ❌
- التصنيف: نظيفة ✅
- السبب: موقف تعليمي طريف, سوء فهم لغوي, مناسب للعائلة

“موقف عائلي” اللهجة: مصرية

النكتة: "مرة واحدة حامل ابنها الصغير ببسألها، ايه اللي في بطنك ده يا ماما قالت له: ده أخوك يا حبيبي، تأني يوم وهو في المدرسة المدرس سألته أنت عندك اخوات؟ قال لها: كان عندي واحد بس ماما أكلته."

التحليل

- النوع: سوء فهم طفولي
- نكات كبار/متزوجين؟ لا ❌ (الحمل مذكور بشكل بريء)
- موت/مرض/حوادث؟ لا ❌ (الطفل فهم خطأ، لا يوجد موت فعلي)
- سخرية قاسية؟ لا ❌
- التصنيف: نظيفة ✅
- السبب: براءة الأطفال في فهم الأمور، كوميديا موقف بريئة

“نكتة البخل اليمني” اللهجة: يمنية

النكتة: "ما زال اليمني لما تقله معك ألف سلف؟ يقولك لو معي مانا هانا، أموت وأعرف أين بيسير لما يكون معه ألف ريال."

التحليل

- النوع: نكتة البخل/الفقر (صورة نمطية اجتماعية)
- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ لا ❌
- العناصر الأساسية:
- اللهجة: يمنية أصيلة ✅
- الشخصية: اليمني (صورة نمطية محلية خفيفة)

“نكتة التكنولوجيا”

اللهجة: مصرية

النكتة: "مرة واحد قفشت جوزها بيخونها على الإنترنت، واما واجهته قالها حرام عليكى هو انتي مشر عارفة إن ده مشر خطي."

التحليل

- النوع: سخرية اجتماعية عن التكنولوجيا
- نكات كبار/متزوجين؟ لا ❌ (ذكر الخيانة لكن بشكل غير جنسي، على الإنترنت فقط)
- موت/مرض/حوادث؟ لا ❌
- سخرية قاسية؟ لا ❌
- التصنيف: نظيفة ✅
- السبب: سخرية من عادات الإنترنت والأعذار الواهية، بدون محتوى جنسي صريح

“نكتة الخليجيين والحلاقة”

اللهجة: مصرية

النكتة: "مرة واحد خليجي دخل على حلاق مصري طلع أصلع، ليش؟ عشان كل ما الحلاق يقوله كويس كده، يرد عليه الخليجي ويقوله: والله ما قمصرت."

التحليل

- النوع: صورة نمطية خفيفة عن الكرم الخليجي
- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ لا ❌
- سخرية قاسية أو عنصرية؟ لا ❌ (صورة نمطية إيجابية عن الكرم، وليست مسيئة)
- التصنيف: نظيفة ✅
- السبب: سخرية لطيفة من سوء التفاهم اللغوي والكرم، مقبولة اجتماعياً

”نكتة الأطفال والآباء”

اللهجة: مصرية

النكتة: ”بابا أنا خارج مع أصحابي النهارده. الأب: انا مش موافق. الابن: ما هو أنت مش خارج معنا.“

التحليل

- النوع: علاقة الآباء والأبناء
- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ لا ❌
- سخرية قاسية؟ لا ❌
- التصنيف: نظيفة ✅
- السبب: فكاهة عائلية خفيفة عن ديناميكيات الأسرة

”نكتة ذمارية”

اللهجة: يمنية

النكتة:

”ذماري اشترى روتي فرنسي، وفي طريقه إلى بيته في القرية سأله أحد جيرانه: إيش اللي بيدك؟ أجاب: هذا روتي فرنسي. سأله الجار: على إيش بتاكله؟ أجاب الذماري: أكله على ملوجة

التحليل

- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ لا ❌
- سخرية قاسية؟ لا ❌
- التصنيف: نظيفة ✅
- السبب: مفارقة ثقافية خفيفة بين طعام أجنبي وطريقة أكل محلية شعبية، بأسلوب لهجي بسيط ومناسب لجميع الأعمار.

الأمثلة المشروحة والحالات الإشكالية

أمثلة نكات سوداء

“الإعدام والموت” اللهجة: مصرية

النكتة: "واحد كسلان أتحكم عليه بالإعدام بدفنه في حفرة, فقالوا له نفسك في إيه قبل ما تموت قال لهم عاوز أشرب قالوا قوم أشرب قال لهم لسه هقوم اردموا يلا."

التحليل

- النوع: فكاهة المشنقة
- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ نعم ✅ (الإعدام والدفن حياً)
- التصنيف: سوداء ●
- السبب: تتناول الموت والإعدام بشكل ساخر، موضوع مأساوي يُقدّم بطريقة فكاهية



“الموت في سياق عائلي” اللهجة: ليبية

النكتة: "واحد خوه مات ويبي يقول لمرت خوه من غير ما يصددها قالها راجلك تزوج عليك قالتله ان شاء الله يخش عليا مكفن قال جيبوه يا شباب الامور طيبة."

التحليل

- النوع: فكاها المشنقة
- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ نعم ✅ (وفاة الأخ، ذكر الكفن)
- التصنيف: سوداء ●
- السبب: السخرية من كيفية إخبار الأخبار السيئة (الموت)، موضوع مأساوي بطريقة فكاهاية

“العنف الشديد” اللهجة: سورية

النكتة: "حمصي محشش قطع رأس أخوه وهو نايم قال اذا قام وما شاف رأسه بيموت ضحك."

التحليل

- النوع: عنف شديد وفكاها سوداء متطرفة
- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ نعم ✅ (قطع الرأس، عنف شديد)
- التصنيف: سوداء ●
- السبب: تصوير عنف دموي قاسٍ بشكل فكاهاي، نكتة سوداء بامتياز

"المرض والمستشفى"

اللهجة: مصرية

النكتة: "واحد بخيل رايج يزور أمه في المستشفى لقي الباب مكتوب عليه ادفع راح لافف وراجع وقال ابقا اروح ازورها في البيت لما ترجع."

التحليل

- النوع: البخل + المرض
- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ نعم ✅ (الأم في المستشفى = مريضة)
- التصنيف: سوداء ●
- السبب: الجمع بين المرض (موضوع حساس) والبخل المبالغ فيه بشكل قاسٍ

"العنصرية القاسية"

اللهجة: سعودية

النكتة: "محشش وهندي طاحوا من الطائرة، المحشش مات والهندي ما مات، سألوا الهندي كيف مات المحشش؟ قال: ياخي هذا نفر همّار، أنا يقول أفتح مظلة، هو يقول: مافي مطر."

التحليل

- النوع: عنصرية + موت
- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ نعم ✅ (السقوط من الطائرة والموت)
- سخرية عنصرية؟ نعم ✅ (استهداف الهنود بصورة نمطية سلبية، واستخدام لفظ "همّار" المسيء)
- التصنيف: سوداء ●
- السبب: تجمع بين الموت والسخرية العنصرية القاسية، مسيئة لمجموعة عرقية

"المرض والمستشفى"

اللهجة: مصرية

النكتة: "واحد بخيل رايج يزور أمه في المستشفى لقي الباب مكتوب عليه ادفع راح لافف وراجع وقال ابقا اروح ازورها في البيت لما ترجع."

التحليل

- النوع: البخل + المرض
- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ نعم ✅ (الأم في المستشفى = مريضة)
- التصنيف: سوداء ●
- السبب: الجمع بين المرض (موضوع حساس) والبخل المبالغ فيه بشكل قاسٍ

"العنصرية القاسية"

اللهجة: سعودية

النكتة: "محشش وهندي طاحوا من الطائرة، المحشش مات والهندي ما مات، سألوا الهندي كيف مات المحشش؟ قال: ياخي هذا نفر همّار، أنا يقول أفتح مظلة، هو يقول: مافي مطر."

التحليل

- النوع: عنصرية + موت
- نكات كبار/متزوجين؟ لا ❌
- موت/مرض/حوادث؟ نعم ✅ (السقوط من الطائرة والموت)
- سخرية عنصرية؟ نعم ✅ (استهداف الهنود بصورة نمطية سلبية، واستخدام لفظ "همّار" المسيء)
- التصنيف: سوداء ●
- السبب: تجمع بين الموت والسخرية العنصرية القاسية، مسيئة لمجموعة عرقية



الأمثلة المشروحة والحالات الإشكالية

أمثلة نكات جنسية

"ليلة الدخلة (تلميحات جنسية)" اللهجة: عراقية

النكتة: "واحد اثول بليلة عرسه ما عنده موضوع يحجي بيه كال
لزوجة اهلع يدرون بيج هنا."

التحليل

- التحليل:
- النوع: نكات كبار/متزوجين
- نكات كبار/متزوجين؟ نعم
- السياق: ليلة العرس - السؤال الغبي للعروس يشير إلى التوتر من الموقف الحميمي
- التصنيف: جنسية ●
- السبب: "ليلة عرسه" هو سياق جنسي بحد ذاته. الموقف المحرج بين العروسين في هذا السياق يحمل معنى جنسياً ضمناً.



“ليلة العرس والمعرفة الجنسية” اللهجة: عراقية

النكتة: "شايب عمره 90 سنة، اتزوج بنية عمره 15 سنة، بليلة الدخلة سألها: بابہ انتي تعرفين شيسوون من يتزوجون؟؟ كالت له: لع، راح ليغاد.. صفن.. كام يحجي ويه نفسه: هاي اشلون ورطة؟؟ البنية متعرف وآني نسيت!!"

التحليل

- النوع: نكات كبار/متزوجين
- نكات كبار/متزوجين؟ نعم
- السياق: سؤال مباشر عن "شيسوون من يتزوجون" = إشارة واضحة للعلاقة الحميمة
- التصنيف: جنسية ●
- السبب: تتحدث بشكل مباشر عن المعرفة الجنسية والعلاقة الحميمة ليلة الزفاف. حتى لو بطريقة كوميدية، المحتوى جنسي واضح.

“ليلة الدخلة (موقف محرج جنسي)” اللهجة: مصرية

النكتة: "في مره واحده رفيعه اوى عاوزه تتجوز ام امها لفت على وسطها جرانيل جرانيل ام واحد شغها عجبته اتجوزوا وفي ليلة الدخلة خلعتها الفستان لقي جرانيل يثيل جرنال يلائي جرنال ام اتصل بحماته والها بنتك على الصفحة الكام."

التحليل

- النوع: نكات كبار/متزوجين
- نكات كبار/متزوجين؟ نعم
- السياق: ليلة الدخلة + "خلعتها الفستان" - موقف حميمي واضح
- التصنيف: جنسية ●
- السبب: تصف موقفاً في ليلة الدخلة عندما يخلع العريس ملابس العروس، وهو موقف جنسي بامتياز. العبارة "بنتك على الصفحة الكام" تشير إلى طبقات الملابس في سياق جنسي.

“المعاني المزدوجة الجنسية (تلميحات صريحة)” اللهجة: مغربية

النكتة: "واحد دري سول الوالد ديالو على الفرق بين الواقع و الحقيقة؟ بدا كيشرح لي،
الدري ما فهم والو!! هو يَكُول ليهِ باه سير كُول لختك: كايِن واحد السعودي باغي يعطي
2 مليون. كَالها لختو، هي تَكُول ليهِ تجيبو ليا نطيك النصف، هو يَكُولها لباه، هو يغوليه
سير كُول هاذ الهضرة لمك خْتَهِيَا، مِثْلا كَالها لها وهي تَكُول ليهِ، سير ا ولدي راك مرضي
جيبو نعطيك النصف. مِثْلا الدري هو ايقولها لباه، هو ايقوليه باه شتي اولدي الحقيقة
هي راه ما كايِن لا سعودي ولا 2 مليون، و الواقع هو راه معانا 2 قح*** فهاذ الدار."

التحليل

- النوع: نكات كبار/متزوجين (محتوى جنسي صريح)
- نكات كبار/متزوجين؟ نعم
- السياق: تتحدث بشكل صريح عن بيع الشرف/الجسد مقابل المال
- الكلمة المحورية: "قح***" (كلمة نابية تشير للنساء في سياق جنسي)
- التصنيف: جنسية ●

“تلميحات جنسية بالتورية (وسائل منع الحمل)” اللهجة: مصرية

النكتة: "ديك كسر بيضه لاقاها فاضيه، دخل للفراخ قالهم مين اللي مركبه لولب."

التحليل

- النوع: نكات كبار/متزوجين (تلميح جنسي بالتورية)
- نكات كبار/متزوجين؟ نعم
- السياق:
- اللولب = وسيلة منع حمل للنساء
- البيضة الفاضية = عدم الإنجاب
- الفراخ = تشبيه بالنساء
- التورية الجنسية واضحة جداً
- التصنيف: جنسية ●
- السبب: إشارة مباشرة لوسيلة منع الحمل (اللولب)

الاعتبارات اللغوية والثقافية وإجراءات التعليق

التنوع اللهجي في النكات العربية

تتميز النكات العربية بتنوع لهجي واسع يمتد من المحيط إلى الخليج.

يواجه المُعلِّقون تحديات في فهم بعض اللهجات، خاصة:

اللهجات الأكثر صعوبة:

- المغاربية (المغرب، الجزائر، تونس، موريتانيا)
- الصومالية
- بعض اللهجات الخليجية الفرعية (عمانية، بحرينية، قطرية)
- اللهجات المحلية الدقيقة (ذمارية، درازية، حضرمية)

استراتيجيات التعامل مع اللهجات غير المألوفة

1. تقنيات الفهم الأولي

أ. التحليل السياقي:

- اقرأ النكتة كاملة عدة مرات
- حاول فهم الموقف العام حتى لو لم تفهم كل كلمة
- ركّز على الكلمات المفتاحية التي قد تحدد التصنيف

ب. استخدام المعرفة المشتركة:

- العديد من الكلمات مشتركة بين اللهجات العربية
- الكلمات الجنسية أو المرضية غالباً واضحة عبر اللهجات
- الأسماء والأفعال الشائعة متشابهة

الاعتبارات اللغوية والثقافية وإجراءات التعليق

التنوع اللهجي في النكات العربية

ج. البحث عن الأنماط:

- نكات المحششين لها بنية متكررة
- نكات البخل لها أنماط مميزة
- نكات الموت غالباً تحتوي كلمات واضحة (مات، توفى، قبر، إلخ)

2. متى تضع علامة [مراجعة]؟

ضع علامة [مراجعة] + [لهجة-صعبة] في الحالات التالية:

✓ حالات المراجعة الإلزامية:

- لا تفهم أكثر من 50% من كلمات النكتة
- اللهجة غير مألوفة تماماً (مثل: صومالية، موريتانية)
- هناك كلمات قد تكون جنسية لكنك غير متأكد من معناها
- التورية معقدة وقد تحمل معنى مزدوج لا تفهمه

✓ حالات المراجعة المستحسنة:

- تشك في أن هناك تلميحاً جنسياً لكنك غير متأكد
- الكلمات قد تكون نابية لكنك لا تعرف اللهجة جيداً
- السياق الثقافي غير واضح (إشارة لحدث أو شخصية لا تعرفها)

حجم العمل والتخطيط الزمني

نموذج التوثيق:

يرجى من كل المشاركين اتباع هذا النموذج عند التعليق على النكته:

- لفئة: [نظيفة/سوداء/جنسية]
- الوسوم: [مراجعة] [لهجة-صعبة]
- التبرير: [اللهجة غير واضحة - فهمت الموقف العام: ...]
- اخترت الفئة بناءً على: [...] مستوى الثقة: [منخفض/متوسط/عالٍ]

ملاحظات:

- الوسوم اختيارية - ضعها فقط عند الحاجة
- التبرير إلزامي للحالات المشكوك فيها فقط
- مستوى الثقة يساعد في المراجعة لاحقاً

حجم العمل والتخطيط الزمني

الوقت المطلوب لكل نكته:

- 2 دقائق لكل نكته (قراءة، فحص، تصنيف، مراجعة)
- متوسط: 2 دقيقة/نكته

حصّة كل مُعلّق:

- 425 نكته لكل مُعلّق

اليوم	عدد النكات	الوقت المتوقع	الراحة
1	85	4 ساعات	4 × 5 دقائق
2	85	4 ساعات	4 × 5 دقائق
3	85	4 ساعات	4 × 5 دقائق
4	85	4 ساعات	4 × 5 دقائق
5	85	4 ساعات	4 × 5 دقائق
الإجمالي	425	20 ساعة	21 فترة

ملخص الدليل

يهدف هذا الدليل إلى توفير إطار عمل موحد ودقيق لتصنيف النكات العربية ضمن ثلاث فئات رئيسية: نظيفة (✓)، سوداء (●)، وجنسية (●). تم تصميم هذا الإطار لبناء نموذج تعلم آلي قادر على فهم وتمييز أنماط الدعابة المختلفة في اللغة العربية بلهجاتها المتنوعة.



النقاط الأساسية:

التصنيف الثلاثي:



- نظيفة (✓): مناسبة لجميع الأعمار، خالية من المحتوى الجنسي والمواضيع المأساوية
- سوداء (●): تتناول مواضيع مأساوية (موت، مرض) أو سخرية قاسية/عنصرية
- جنسية (●): تحتوي على محتوى أو تلميحات جنسية

أولويات التصنيف:



- جنسية < سوداء < نظيفة
- عند التداخل، اختر الفئة الأكثر تقييداً

الاعتبارات الثقافية:



- احترام التنوع اللهجي في العالم العربي
- فهم السياق الثقافي لكل منطقة
- الموضوعية والحياد في التصنيف

إجراءات التعليق:



- قراءة متأنية (مرتين على الأقل)
- اتباع شجرة القرار
- توثيق الحالات المشكوك فيها

الأهمية البحثية للمشروع

هذا المشروع يساهم في:

1. تطوير تقنيات معالجة اللغة العربية الطبيعية (NLP)

- بناء نماذج قادرة على فهم الفكاهة العربية
- تحسين قدرة الأنظمة الذكية على فهم السياقات الفكاهية
- المساهمة في أبحاث الحوسبة اللسانية العربية

2. الدراسات الثقافية واللغوية

- توثيق أنماط الفكاهة عبر اللهجات العربية المختلفة
- دراسة الخصائص اللغوية والثقافية للدعابة العربية
- فهم التنوع الثقافي في استخدام الفكاهة

3. التطبيقات العملية

- تطوير أنظمة ترشيح المحتوى الذكية
- تحسين خوارزميات التوصية بالمحتوى المناسب
- بناء مساعدات رقمية أكثر ذكاءً في فهم السياق الكوميدي

رسالة ختامية للمُعلقين

عزيزي المُعلق،

عملك في هذا المشروع ذو قيمة علمية عالية. أنت تساهم في بناء أول مجموعة بيانات مُصنّفة بدقة للنكات العربية، والتي ستُستخدم في تدريب نماذج ذكاء اصطناعي قادرة على فهم الفكاهة العربية بتنوعها الثري.

تذكّر دائماً:

1. ✓ الجودة أهم من السرعة - خذ وقتك في فهم كل نكتة وتصنيفها بدقة
2. ✓ الموضوعية مطلوبة - اترك آراءك الشخصية جانباً وركّز على المحتوى الفعلي
3. ✓ الاتساق مفتاح النجاح - حافظ على نفس المعايير عبر جميع النكات التي تُصنّفها
4. ✓ لا تتردد في طلب المساعدة - ضع علامة [مراجعة] عندما تشك، فهذا أفضل من تخمين خاطئ
5. ✓ أنت جزء من بحث علمي - عملك سيساهم في تطوير تكنولوجيا معالجة اللغة العربية

شكر وتقدير

نشكرك على مساهمتك القيّمة في هذا المشروع البحثي. التزامك بالدقة والموضوعية سيضمن جودة البيانات وبالتالي نجاح النموذج اللغوي الذي سيبنى عليها.

الإصدار: 1.0

تاريخ آخر تحديث: ديسمبر 2025

<https://osamah-s7s.github.io/>



osamah0alini@gmail.com



+92-333-419-4029

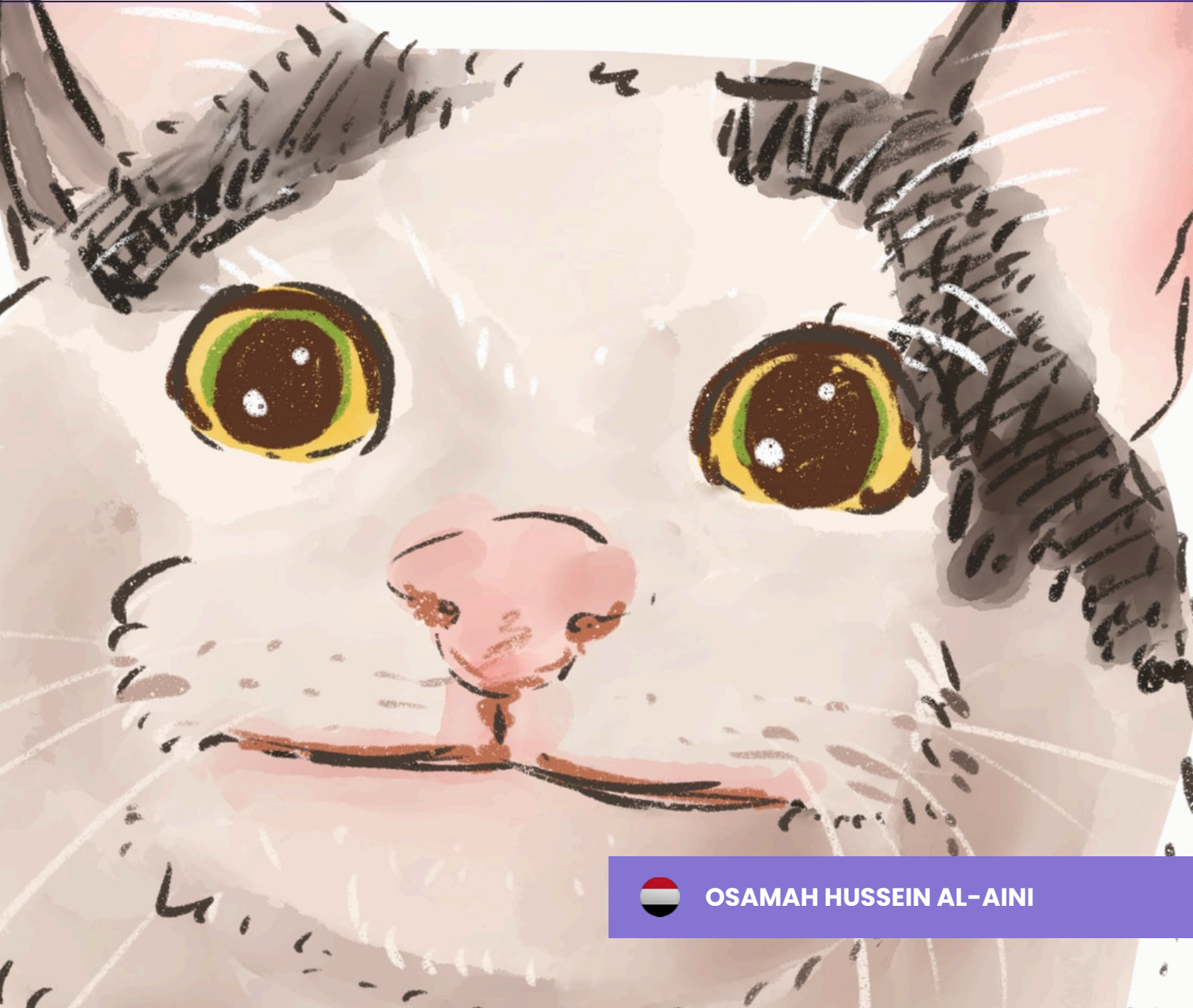


Gulberg III, Street 47, Lahore,
Pakistan



معلومات الاتصال والدعم

إذا كانت لديك أي أسئلة أو استفسارات حول
التصنيف:



OSAMAH HUSSEIN AL-AINI