

مهارات بحثية 1 التكليف 5

دراسة أهداف وفروض البحث

Prepared by: **Abobakr Mohammed Baras**
ID: **253002656**

Nov 2025





Enrichment of Arabic WordNet Using Machine Translation and Transformers

Link: <https://aclanthology.org/2023.icnlp-1.36.pdf>

لمحة عامة

Chapter 01

هدف الدراسة: تحسين تغطية وجودة الشبكة الدلالية العربية من خلال الاستفادة من النسخة الحديثة لـ Princeton WordNet وتقنيات التعلم العميق.

المنهجية: ترجمة متعددة المصادر: استخدام عدة أنظمة ترجمة لتحويل بيانات PWN والبيانات المستخرجة من الويب إلى العربية. تحسين الترجمة بالتعلم العميق: توظيف نماذج Transformer لتصفية وتحسين المخرجات المترجمة. تطوير بوابة ويب: تمكّن المستخدمين والخبراء من عرض المقترحات والتحقق من صحة العلاقات الدلالية.

النتائج:

تحقق دقة 75.4% في التحقق اليدوي لعينة عشوائية من 1000 سينسيت.

مع هذه الدقة، يمكن الحصول على حوالي 68,000 سينسيت صحيحة من أصل 90,127 مرشحاً للإثراء.

الهدف العام: تحسين غطاء وجودة الشبكة الدلالية للغة العربية عبر استغلال PWN وتقنيات التعلم العميق.

أهداف فرعية:

اضافة مرادفات (Synsets) عربية جديدة مطابقة لبنية PWN.

دمج تعريفات وأمثلة عربية مُستخرجة/مترجمة لكل Synset لرفع قابلية الاستخدام.

استعمال عدة محركات ترجمة ثم تنقية النتائج بواسطة Transformers (مثل MT5 / AraBERT) لتحسين الدقة السياقية.

توفير منصة ويب للتحقق البشري والتصديق على المرشحات قبل إدراجها.

مقاييس الأداء المذكورة في الورقة:

عدد المرشحات المُولدة: Synsets 90,127 (مرشحة).

تقييم تجريبي على عيّنة عشوائية Synset 1,000: دقة التحقق اليدوي الإجمالية $\approx 75.4\%$ → تقدير Synsets 68,000 صحيحة من 90,127 مرشحة (تقديري بحسب الدقة).

هل الورقة تذكر فروضاً صريحة؟

لا توجد فروض صريحة ومكتوبة بصيغة اختبارية في النص؛ الورقة تعرض أهدافاً ومنهجية و نتائج تجريبية لكنها لا تصيغ فروضاً على شكل فروض صفرية بصيغة العدم أو فروض مذكورة بصيغة الإثبات.

(المصدر: الملخص و مقدمة الورقة).

الفرض المقترح 1 — رفع الغطاء (Coverage)

- الفرض الصفري: إضافة المنهجية المقترحة لا تؤدي إلى زيادة في عدد Synsets المعتمدة في AWN.
 - الفرض البديل: إضافة المنهجية المقترحة تؤدي إلى زيادة معنوية في عدد Synsets المعتمدة في AWN.
- (مبرّر: الورقة تذكر توليد 90,127 مرشحًا وإمكانية قبول ~68,000 بناءً على دقة 75.4%).

الفرض المقترح 2 — جودة التعريفات والأمثلة (Accuracy)

- الفرض الصفري: المنهجية لا تحقق دقة تحقق يختلف عن الدقة الموجودة حالياً في AWN
 - الفرض البديل: المنهجية تحقق تحسناً معنوياً في دقة التعريفات والأمثلة
- (مبرّر: الورقة تقيس دقة التعريفات والكلمات والأمثلة في عينة 1,000).

الفرض المقترح 3 — فاعلية استخدام محركات ترجمة + Transformers

- الفرض الصفري: استخدام سلسلة (ترجمة متعددة) متبوعاً بتنقية Transformers لا يحسّن جودة الترجمات المختارة مقارنة باستخدام مترجم واحد.
 - الفرض البديل: استخدام سلسلة (ترجمة متعددة) ثم تنقية بـ Transformers يحسّن جودة الترجمات والسياق مقارنة بمترجم واحد.
- (مبرّر: الورقة تقارن Google/MT5/Reverso وتستخدم Back-and-Forth + AraBERT/MT5 لاختيار الأفضل).



مجموعة جامعة
الجميع الذكية
للتعليم القابضة

MIDOCEAN
UNIVERSITY



Thank You

