



MGB
CHALLENGE

MGB-5 (Multi-Genre Broadcast) 2019

تحدي البث متعدد الأنواع (MGB-5)

ركزت MGB-5 على التعرف على الكلام العربي المغربي؛ وتحديد اللهجات العربية المختلفة في مقاطع فيديو يوتيوب من ١٧ دولة ناطقة بالعربية.

حازت (الشركة الهندسية لتطوير النظم الرقمية – جامعة RDI-CU) على المركز الأول

تم تقسيم التحدي إلى مهمتين. كانت المهمة الأولى هي التعرف على ١٤ ساعة من مقاطع الفيديو التي تحتوي على كلام مغربي وتحويلها إلى نص مكتوب.

كان محتوى مقاطع الفيديو متنوع وشامل لمختلف الاهتمامات مثل: الكوميديا والطبخ والأسرة / الأطفال والأزياء والدراما والرياضة والعلوم (TEDx).

المهمة الثانية كانت التعرف على اللهجة العربية الموجودة في مقاطع فيديو مختلفة من ١٧ دولة عربية. تم تقسيم بيانات تعريف اللهجة إلى ثلاث فئات فرعية بناءً على مدة المقطع: قصيرة (أقل من خمس ثوانٍ) ومتوسطة (٥-٢٠ ثانية) وطويلة (> ٢٠ ثانية).

تم الاعلان عن فوز RDI-CU بالمهمة الأولى في التحدي، حيث حصلنا على أدنى معدلات الخطأ في مهمة تحويل الكلام إلى نص، ولم تشارك RDI-CU في المهمة الثانية.

| | MGB5 WER PER TRANSCRIBER | | | | MGB5 | |
|----------|--------------------------|------|------|------|--------|--------|
| | WER1 | WER2 | WER3 | WER4 | AV-WER | MR-WER |
| RDI-CU | 59.1 | 58.0 | 60.1 | 60.1 | 59.4 | 37.6 |
| DARTS | 62.3 | 62.2 | 62.9 | 63.6 | 62.7 | 41.8 |
| Baseline | 66.8 | 66.9 | 67.2 | 67.6 | 67.1 | 48.4 |
| ZXIAT | 67.3 | 67.2 | 67.7 | 67.8 | 67.5 | 49.25 |

رابط التحدي: <http://www.mgb-challenge.org/MGB-5.html>

رابط ورقة البحث والتي حددت النتائج: http://arabicspeech.qcri.com/arabicspeech-portal-resources/papaers/MGB5_ASRU2019.pdf