

ML_DS_TAKEHOME ASSIGNMENT EVALUATION REPORT

| Metric | Value | Explanation |
|----------------------------|--------|--|
| 1. Matched Fields | 86 | Number of fields that were correctly predicted (i.e., they match the ground truth exactly). |
| 2. Predicted Fields | 129 | Total number of fields the OCR system attempted to extract. |
| 3. Ground Truth | 119 | Total number of correct fields that should have been extracted (the reference answer). |
| 4. Precision | 0.6667 | Out of all predicted fields, approximately 66.7% were correct. Formula: Matched / Predicted = 86 / 129. |
| 5. Recall | 0.7227 | Out of all actual fields in the ground truth, approximate 72.3% were correctly found. Formula: Matched / Ground Truth = 86 / 119. |
| 6. F1-Score | 0.6935 | Harmonic mean of precision and recall. It's a balanced metric for overall quality. Formula: $2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$. |

Interpretation

- The **precision** of 66.7% suggests that there were quite a few false positives extra fields the model predicted that didn't exist in the ground truth.
- The **recall** of 72.3% means the OCR captured most of the actual fields but still missed some.
- An **F1-score** of 69.4% gives a single score balancing both errors: missed fields (false negatives) and extra wrong fields (false positives).

The Extraction Process can still be fine-tune for mode accuracy.