**Supplementary Table 1:** Most frequent SNOMED-3 codes and their equivalent SNOMED-CT (Clinical Terms) terms, along with the number of cases for each class. Additionally, the table includes the best performance achieved by the models, the difference in terms of F1 score (Diff.) between the two models (major differences in orange), the mean text length (extreme values ​​in orange) and the model that performed better. M, Morphologic abnormality; D, Disorder; F, Finding; Q, Qualifier value.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **SNOMED-3** | **SNOMED-CT (concept type)** | **N° Cases**  | **Best value** | **Diff.** | **Mean text length**  | **Best model**  |
| M-80812 | Squamous cell carcinoma in situ (M) | 20 | 1 | 0.20 | 64 | SVM |
| M-87500 | Intradermal nevus (M) | 80 | 1 | 0.01 | 61 | SVM |
| M-88100 | Benign fibroma (M) | 11 | 1 | 0 | 61 | SVM |
| M-67032 | Mild dysplasia (M) | 641 | 0.98 | 0.02 | 79 | SVM |
| M-87600 | Compound nevus (M) | 141 | 0.98 | 0 | 66 | SVM |
| M-88500 | Lipoma (M) | 30 | 0.98 | 0.05 | 22 | SVM |
| M-33410 | Epidermoid cyst (M) | 14 | 0.97 | 0.05 | 72 | SVM |
| M-80520 | Squamous cell papilloma (M) | 16 | 0.97 | 0.05 | 47 | SVM |
| M-72200 | Lymphoid hyperplasia (M) | 42 | 0.96 | 0.75 | 82 | SVM |
| M-72850 | Actinic keratosis (M) | 24 | 0.96 | 0.17 | 70 | SVM |
| M-79310 | Proliferative endometrium (M) | 17 | 0.94 | 0.02 | 83 | SVM |
| M-87270 | Dysplastic nevus (M) | 43 | 0.94 | 0.01 | 83 | SVM |
| M-80703 | Squamous cell carcinoma (M) | 108 | 0.93 | 0.19 | 198 | SVM |
| M-87433 | Malignant melanoma (M) | 22 | 0.93 | 0 | 550 | SVM |
| M-85003 | Infiltrating duct carcinoma (M) | 79 | 0.92 | 0.70 | 392 | SVM |
| M-88900 | Leiomyoma (M) | 7 | 0.92 | 0 | 109 | SVM |
| M-09010 | Inadequate (Q) | 35 | 0.91 | 0.03 | 124 | SVM |
| M-72750 | Seborrheic keratosis (M) | 41 | 0.91 | 0.67 | 58 | SVM |
| M-82130 | Sessile serrated adenoma (M) | 36 | 0.91 | 0 | 79 | SVM |
| M-87400 | Junctional nevus (M) | 46 | 0.91 | 0 | 68 | SVM |
| M-79320 | Secretory endometrium (M) | 19 | 0.9 | 0.80 | 89 | SVM |
| M-58000 | Atrophic (M) | 326 | 0.87 | 0.37 | 183 | SVM |
| M-67033 | Severe dysplasia (M) | 20 | 0.86 | 0.29 | 130 | SVM |
| M-54701 | Focal infarct (M) | 11 | 0.85 | 0.05 | 102 | SVM |
| M-80943 | Basosquamous carcinoma (M) | 11 | 0.79 | 0.04 | 129 | SVM |
| M-80106 | Carcinoma, metastatic (M) | 39 | 0.75 | 0.02 | 181 | SVM |
| M-81406 | Metastatic adenocarcinoma (M) | 58 | 0.73 | 0 | 137 | SVM |
| M-40000 | Mild inflammation (M) | 11 | 0.67 | 0.31 | 157 | SVM |
| M-09410 | Negative for residual tumor (F) | 28 | 0.26 | 0.03 | 98 | SVM |
| M-81302 | Papillary transitional cell carcinoma, non-invasive (M) | 104 | 1 | 0.01 | 176 | LSTM |
| M-80903 | Basal cell carcinoma (M) | 437 | 0.98 | 0.05 | 72 | LSTM |
| M-81203 | Transitional cell carcinoma (M) | 44 | 0.98 | 0.05 | 194 | LSTM |
| M-26520 | Odontogenic cyst (M) | 15 | 0.97 | 0.06 | 65 | LSTM |
| M-67060 | Atypia suspicious for malignancy (M) | 100 | 0.97 | 0.04 | 210 | LSTM |
| M-76820 | Inflammatory polyp (M) | 19 | 0.97 | 0.38 | 61 | LSTM |
| M-72400 | Glandular and stromal hyperplasia (M) | 345 | 0.96 | 0.17 | 143 | LSTM |
| M-76800 | Polyp (M) | 35 | 0.95 | 0.03 | 87 | LSTM |
| M-81403 | Adenocarcinoma (M) | 369 | 0.95 | 0 | 288 | LSTM |
| M-90100 | Benign fibroadenoma (M) | 20 | 0.95 | 0.03 | 105 | LSTM |
| M-73320 | Intestinal metaplasia (M) | 395 | 0.94 | 0.03 | 252 | LSTM |
| M-79370 | Atrophic endometrium (M) | 13 | 0.91 | 0.32 | 108 | LSTM |
| M-80003 | Malignant neoplasm (M) | 156 | 0.91 | 0.41 | 269 | LSTM |
| M-41000 | Acute inflammation (M) | 519 | 0.9 | 0.64 | 105 | LSTM |
| D5-74100 | Celiac disease (D) | 21 | 0.89 | 0.03 | 148 | LSTM |
| M-74000 | Dysplasia (M) | 26 | 0.87 | 0.33 | 112 | LSTM |
| M-72042 | Hyperplastic polyp (M) | 37 | 0.84 | 0.18 | 41 | LSTM |
| M-43000 | Chronic inflammation (M) | 1484 | 0.82 | 0.69 | 151 | LSTM |
| M-33400 | Cyst (M) | 6 | 0.67 | 0.41 | 97 | LSTM |
| M-54700 | Infarct (M) | 12 | 0.67 | 0.40 | 55 | LSTM |
| M-00110 | Normal tissue morphology (F) | 230 | 0.52 | 0.04 | 87 | LSTM |