**A Dataset details**

**1 Tumor Abbreviation, Coding and other information**

See files named common\_samples.csv, tumor\_id.csv, tumor\_TSS.csv.

**2 Training, Validation and Testing dataset**

See files named train.txt, validation.txt, test.txt.

**3 Tumor samples information**

|  |  |  |
| --- | --- | --- |
| Tumor type | Tumor id | Tumor number |
| Testicular Germ Cell Tumors | 0 | 122 |
| Head and Neck squamous cell carcinoma | 1 | 338 |
| Breast invasive carcinoma | 2 | 626 |
| Bladder Urothelial Carcinoma | 3 | 419 |
| Lymphoid Neoplasm Diffuse Large B-cell Lymphoma | 4 | 33 |
| Cervical squamous cell carcinoma and endocervical adenocarcinoma | 5 | 168 |
| Liver hepatocellular carcinoma | 6 | 178 |
| Rectum adenocarcinoma | 7 | 73 |
| Kidney renal clear cell carcinoma | 8 | 323 |
| Stomach adenocarcinoma | 9 | 293 |
| Colon adenocarcinoma | 10 | 239 |
| Sarcoma | 11 | 219 |
| Esophageal carcinoma | 12 | 124 |
| Pheochromocytoma and Paraganglioma | 13 | 82 |
| Glioblastoma multiforme | 14 | 32 |
| Prostate adenocarcinoma | 15 | 349 |
| Kidney Chromophobe | 16 | 62 |
| Uterine Corpus Endometrial Carcinoma | 17 | 293 |
| Skin Cutaneous Melanoma | 18 | 348 |
| Kidney renal papillary cell carcinoma | 19 | 249 |
| Uveal Melanoma | 20 | 12 |
| Pancreatic adenocarcinoma | 21 | 97 |
| Lung adenocarcinoma | 22 | 53 |
| Uterine Carcinosarcoma | 23 | 48 |
| Cholangiocarcinoma | 24 | 30 |
| Thymoma | 25 | 113 |
| Lung squamous cell carcinoma | 26 | 26 |
| Chronic Myelogenous Leukemia | 27 | 225 |
| Thyroid carcinoma | 28 | 368 |
| Mesothelioma | 29 | 61 |
| Brain Lower Grade Glioma | 30 | 432 |
| Ovarian serous cystadenocarcinoma | 31 | 53 |
| Adrenocortical carcinoma | 32 | 45 |