

```
In [1]: import os, sys
import pandas as pd
from pathlib import Path

if os.path.exists(os.path.abspath('../src')):
    sys.path.append(os.path.abspath('../src'))

from feat_build import main
from model_build import train
```

Intel MKL WARNING: Support of Intel(R) Streaming SIMD Extensions 4.2 (Intel(R) SSE4.2) enabled only processors has been deprecated. Intel oneAPI Math Kernel Library 2025.0 will require Intel(R) Advanced Vector Extensions (Intel(R) AVX) instructions.

Intel MKL WARNING: Support of Intel(R) Streaming SIMD Extensions 4.2 (Intel(R) SSE4.2) enabled only processors has been deprecated. Intel oneAPI Math Kernel Library 2025.0 will require Intel(R) Advanced Vector Extensions (Intel(R) AVX) instructions.

```
In [2]: inv_dir = Path(os.getcwd())
proj_dir = inv_dir.parent

sample_guids_parquet = 'sample_guid_10000_china_us.parquet'

if 'feat.parquet' not in os.listdir(inv_dir / 'out'):
    main.generate_features(sample_guids_parquet, inv_dir)
```

web_usage exists...

sw_usage exists...

temp exists...

cpu_util exists...

power exists...

Loading raw data...

Processing raw data...

Processing sysinfo data...

```
/Volumes/XtremeSSD/CapstoneIntelDP/Data_Intel_Globus/REPO_CLONE/src/feat_build/sysinfo_process.py:25: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
sysinfo.replace({'n/a': 'Unknown', 'N/A': 'Unknown'}, inplace=True)  
/Volumes/XtremeSSD/CapstoneIntelDP/Data_Intel_Globus/REPO_CLONE/src/feat_build/sysinfo_process.py:28: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
sysinfo['cpu_suffix'] = sysinfo['cpu_suffix'].apply(lambda x: x if 'Core' in x else 'Other')  
/Volumes/XtremeSSD/CapstoneIntelDP/Data_Intel_Globus/REPO_CLONE/src/feat_build/sysinfo_process.py:31: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
sysinfo['screen_size_category'] = sysinfo['screen_size_category'].str.replace('x', '').replace(screen_size_mapping)  
/Volumes/XtremeSSD/CapstoneIntelDP/Data_Intel_Globus/REPO_CLONE/src/feat_build/sysinfo_process.py:88: SettingWithCopyWarning:  
A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy

```
chastype['portable'] = chastype['chassistype'].isin(portable).astype(int)
```

Joining and standardizing data...

Writing features to disk...

Feature processing complete! feat.parquet saved to /Volumes/XtremeSSD/CapstoneIntelDP/Data_Intel_Globus/REPO_CLONE/inv/out

```
In [5]: feat = pd.read_parquet(os.path.join('out', 'feat.parquet'))
```

```
In [6]: test_mse, feat_dict, r2 = train.train(feat)
```

Test MSE: 10.36

```
In [7]: r2
```

```
Out[7]: 0.6521609211717979
```

```
In [8]: feat_dict
```

```
Out[8]: {'sw_category_Development & Programming (IDEs, Text Editors, Version
```

Control)': -9966065216.910608,
'sw_category_Gaming (Casual, Online & Offline)': -29836274931.00190
7,
'sw_category_Multimedia Editing (Audio & Video)': -20226995207.5537
1,
'sw_category_Other': -1346172734820.7742,
'sw_category_Productivity & Office': -85037204320.80908,
'sw_category_Security & Network': -17177768614.561829,
'sw_category_Simulation & Virtual Reality': -19604508742.090942,
'sw_category_System & Utilities': -14228456882.803772,
'sw_category_Web Browsers & Communication': -212720449044.3955,
'sw_event_name_DC_DISPLAY_OFF': 1349139339109.742,
'sw_event_name_DC_DISPLAY_ON': 214231547217.96832,
'temp_avg': 2.925811767578125,
'web_parent_category_content creation': -1817320873144.1282,
'web_parent_category_education': -4129294082367.8,
'web_parent_category_entertainment': 6862452373127.241,
'web_parent_category_finance': 545299703771.33685,
'web_parent_category_games': 882057609166.5731,
'web_parent_category_mail': -711219997789.2192,
'web_parent_category_news': -557484329999.2589,
'web_parent_category_other': 4407660593438.249,
'web_parent_category_private': 1437430272688.5515,
'web_parent_category_productivity': 5426990092194.151,
'web_parent_category_recreation': -825986620364.2361,
'web_parent_category_reference': 2038625773319.1523,
'web_parent_category_search': 4459469616400.649,
'web_parent_category_shopping': 2218860966326.6494,
'web_parent_category_social': -799737587124.9266,
'web_sub_category_banking and accounting': -545299703771.3472,
'web_sub_category_communication': 351929587264.1828,
'web_sub_category_communication - live': 204903338484.66556,
'web_sub_category_crm': -4829416469.825541,
'web_sub_category_education': 4129294082367.8647,
'web_sub_category_mail': 711219997789.2335,
'web_sub_category_music / audio streaming': -321354653373.5369,
'web_sub_category_news': 557484329999.2365,
'web_sub_category_other': -4773351155556.201,
'web_sub_category_photo edit/creation': 630231061632.3832,
'web_sub_category_presentations': -599304499946.2528,
'web_sub_category_private': -1437430272688.601,
'web_sub_category_programming': -183000481070.58304,
'web_sub_category_project management': -439796932751.9755,
'web_sub_category_reference': -2038625773319.1658,
'web_sub_category_search': -4459469616400.645,
'web_sub_category_shopping': -2218860966326.6475,
'web_sub_category_social network': 684237823803.1577,
'web_sub_category_spreadsheets': -976793478599.3274,
'web_sub_category_travel': 825986620364.2257,
'web_sub_category_unclassified': -4407660593438.254,
'web_sub_category_video games': -30138114284.39703,
'web_sub_category_video streaming': -6394819626039.075,

'web_sub_category_video/audio edit/creation': 58576873980.605156,
'web_sub_category_web design / development': 1702603327920.1863,
'web_sub_category_word processing': -3417352029778.6226,
'cpu_norm_usage': 0.7783203125,
'nrs': 0.065826416015625,
'countryname_normalized_Argentina': -84018657154.99219,
'countryname_normalized_Australia': -18304936210.20777,
'countryname_normalized_Austria': -1224895204362.3906,
'countryname_normalized_Bangladesh': -1173083665900.4712,
'countryname_normalized_Belgium': 1274033233806.2168,
'countryname_normalized_Brazil': 1450588784513.3796,
'countryname_normalized_Canada': -2261052034863.497,
'countryname_normalized_Chile': -819444009513.966,
'countryname_normalized_China': 530714007930.9835,
'countryname_normalized_Colombia': -478336013005.29535,
'countryname_normalized_Czech Republic': -119772342087.59042,
'countryname_normalized_Denmark': -684105749539.4712,
'countryname_normalized_Ecuador': 189172634799.15405,
'countryname_normalized_Egypt': 975916765686.6372,
'countryname_normalized_France': 123642800112.82913,
'countryname_normalized_Germany': 94912470969.05843,
'countryname_normalized_Greece': -380434281769.3116,
'countryname_normalized_Hong Kong': -172246742326.72037,
'countryname_normalized_Hungary': -57295387940.19352,
'countryname_normalized_India': 314927446624.8663,
'countryname_normalized_Indonesia': 85012141453.4807,
'countryname_normalized_Israel': -36342240177.988045,
'countryname_normalized_Italy': 5067674740.399342,
'countryname_normalized_Japan': 233362282939.1819,
'countryname_normalized_Korea, Republic of': 73827078389.28107,
'countryname_normalized_Malaysia': -61909205789.74685,
'countryname_normalized_Mexico': -26292838575.84061,
'countryname_normalized_Netherlands': -13410375215.6766,
'countryname_normalized_Norway': -19018213533.22139,
'countryname_normalized_Other': 15303601925.75865,
'countryname_normalized_Pakistan': 43590606759.41302,
'countryname_normalized_Peru': -3530502594.649826,
'countryname_normalized_Philippines': 6682830028.645107,
'countryname_normalized_Poland': 1794915338.7880936,
'countryname_normalized_Portugal': -3124117308.681571,
'countryname_normalized_Romania': -16389769162.699245,
'countryname_normalized_Russian Federation': -16420878374.830858,
'countryname_normalized_Saudi Arabia': -22603665433.3293,
'countryname_normalized_Singapore': 4399397855.312575,
'countryname_normalized_South Africa': 20287240773.9308,
'countryname_normalized_Spain': 9343071166.398762,
'countryname_normalized_Sweden': 11708616437.698494,
'countryname_normalized_Switzerland': -7365558389.339016,
'countryname_normalized_Taiwan, Province of China': 1845095969.89241
43,
'countryname_normalized_Thailand': -79674804.90619016,
'countryname_normalized_Turkey': 7740251192.782366,

'countryname_normalized_Ukraine': -235655911.85855746,
'countryname_normalized_United Arab Emirates': -1934019892.8787038,
'countryname_normalized_United Kingdom of Great Britain and Northern Ireland': 2142005642.8077962,
'countryname_normalized_United States of America': 530714007930.2654
4,
'countryname_normalized_Viet Nam': 387942388.1900336,
'modelvendor_normalized_AZW': -5488356.440397374,
'modelvendor_normalized_Acer': 4254385278450.978,
'modelvendor_normalized_Alienware': 4254385278450.128,
'modelvendor_normalized_Apple': 4254385278449.281,
'modelvendor_normalized_Asus': 4254385278451.0757,
'modelvendor_normalized_Biostar': 222145314.87329102,
'modelvendor_normalized_Casper Bilgisayar': -361442001.5390625,
'modelvendor_normalized_Clevo': 4254385278436.1255,
'modelvendor_normalized_Colorful Technology And Development Co.,LT
D': -0.00390625,
'modelvendor_normalized_Dell': 4254385278450.4297,
'modelvendor_normalized_ECS': 0.00048828125,
'modelvendor_normalized_Fujitsu': 4254385278451.4346,
'modelvendor_normalized_GPD': 4254385278458.554,
'modelvendor_normalized_Gateway': -0.001953125,
'modelvendor_normalized_Gigabyte': 4254385278450.2695,
'modelvendor_normalized_HANSUNG COMPUTER': 0.0,
'modelvendor_normalized_HP': 4254385278451.1226,
'modelvendor_normalized_HUAWEI': 4254385278451.037,
'modelvendor_normalized_Hasee': 4254385278452.233,
'modelvendor_normalized_Intel': 4254385278454.1523,
'modelvendor_normalized_LG': 4254385278448.69,
'modelvendor_normalized_Lenovo': 4254385278451.738,
'modelvendor_normalized_MECHREVO': 4254385278457.424,
'modelvendor_normalized_MONSTER': 0.0,
'modelvendor_normalized_MSI': 4254385278451.17,
'modelvendor_normalized_Medion': 4254385278452.711,
'modelvendor_normalized_Microsoft Corporation': 0.0,
'modelvendor_normalized_Mouse Computer': 4254385278447.1187,
'modelvendor_normalized_NEC': 0.0,
'modelvendor_normalized_Notebook': 4254385278452.662,
'modelvendor_normalized_O.E.M': 4254385278452.3994,
'modelvendor_normalized_Other': 4254385278452.237,
'modelvendor_normalized_PCWARE': 0.0,
'modelvendor_normalized_Packard Bell': 0.0,
'modelvendor_normalized_Panasonic Corporation': 4254385278452.913,
'modelvendor_normalized_Positivo': 0.0,
'modelvendor_normalized_Razer': 4254385278450.2964,
'modelvendor_normalized_Samsung': 4254385278450.1675,
'modelvendor_normalized_Shinelon Computer': 4254385278450.1836,
'modelvendor_normalized_Sony': 0.0,
'modelvendor_normalized_Supermicro': 0.0,
'modelvendor_normalized_THUNDEROBOT': 4254385278454.5713,
'modelvendor_normalized_TIMI': 4254385278452.7554,
'modelvendor_normalized_Thirdwave Corporation': 0.0,

'modelvendor_normalized_Timi': 4254385278449.6865,
'modelvendor_normalized_Toshiba': 4254385278450.7524,
'modelvendor_normalized_Unknown': 4254385278450.8115,
'modelvendor_normalized_Wortmann_AG': 4254385278451.166,
'os_Unknown': 0.0,
'os_Win Server': 0.0,
'os_Win10': 2791596045550.4385,
'os_Win11': 2791596045550.2505,
'os_Win7': 0.0,
'os_Win8': 0.0,
'os_Win8.1': 0.0,
'graphicsmanuf_AMD': -117486084908.999,
'graphicsmanuf_Intel': -117486084908.21855,
'graphicsmanuf_Nvidia': -117486084907.98802,
'graphicsmanuf_Other': -117486084907.53673,
'cpu_family_Atom': 0.0,
'cpu_family_Core M': -289345462197.4429,
'cpu_family_Core i3': -289345462198.89417,
'cpu_family_Core i5': -289345462200.5051,
'cpu_family_Core i7': -289345462200.0783,
'cpu_family_Core i9': -289345462199.76636,
'cpu_family_Core2': 0.0,
'cpu_family_Other': -289345462180.3648,
'cpu_family_Pentium/Celeron': -289345462199.89624,
'cpu_family_Xeon': -289345462200.595,
'cpu_suffix_Core-C': 0.0,
'cpu_suffix_Core-E': 0.0,
'cpu_suffix_Core-G': -1844573966576.3765,
'cpu_suffix_Core-H': -1844573966574.2075,
'cpu_suffix_Core-K': -1844573966568.954,
'cpu_suffix_Core-M': 0.0,
'cpu_suffix_Core-No suffix': -1844573966569.5618,
'cpu_suffix_Core-P': 0.0,
'cpu_suffix_Core-R': 0.0,
'cpu_suffix_Core-S': 0.0,
'cpu_suffix_Core-T': 0.0,
'cpu_suffix_Core-U': -1844573966575.718,
'cpu_suffix_Core-X': 0.0,
'cpu_suffix_Core-Y': -1844573966578.2192,
'cpu_suffix_Core2 Duo': 0.0,
'cpu_suffix_Core2 Extreme': 0.0,
'cpu_suffix_Core2 Quad': 0.0,
'cpu_suffix_Core2 Solo': 0.0,
'cpu_suffix_Other': -1844573966574.7068,
'persona_Casual Gamer': 94435423627.53473,
'persona_Casual User': 94435423627.80434,
'persona_Communication': 94435423626.7103,
'persona_Content Creator/IT': 94435423627.26767,
'persona_Entertainment': 94435423626.88144,
'persona_File & Network Sharer': 94435423626.99496,
'persona_Gamer': 94435423627.87433,
'persona_Office/Productivity': 94435423627.00761,

```
'persona_Unknown': 94435423626.66452,  
'persona_Web User': 94435423627.43817,  
'persona_Win Store App User': 94435423628.3978,  
'age_category': 0.4107666015625,  
'ram': 276.9571245908737,  
'#ofcores': 1.8390960693359375,  
'screen_size_category': 2.096099853515625,  
'day_of_week': -0.023681640625,  
'month_of_year': -0.016998291015625}
```

In []: