Konstantin Privalov

+7-950-220-80-90 | Location: Saint Petersburg, ready to relocate | karkadevich@yahoo.co.uk | kaggle.com/beifaa | github.com/beifa

Freelance

upwork | *PyTorch, Opencv, ONNX, Docker,streamlit, wandb, BoofCV, supervise.ly ...* 2021-04 ...2021-08

- By photo from the camera, detects QR codes, recognizes(decoder QR coders)them and get a list of QR codes and show on image. Labeled data, fasterrcnn box/mask, image processing(super resolution, perspective transform and other) to improve QR code recognition, presentation for client by dash(plotly), deploy onnx model to Jetson Nano and more...

Projects(Kaggle)

Rainforest Connection Species Audio Detection, 73 place | *Librosa, PyTorch, Optuna, Pandas..*

* In this competition, you'll automate the detection of bird and frog species in tropical soundscape recordings.
* I transform flac to mel spectrogram(use Fast Fourier Transform, Mel scale)and make image
* Use stratified fold 5 splits,i invented make fold by bin different frequency flacs, metric label-ranking average precision,i find best heads layers and albumentations
* Make Ensemble (efficientne tb0-b4, mobilenet) with different params by mel spectogram, used meta data from flacs
* ban multi accounts...

ASHRAE-Great Energy Predictor III, Top 11 % | *KMeas, LGBM, Catboost, Ridge, Lasso, Pandas, Numpy*

* In this competition, develop accurate models of metered building energy usage
* Used Hyperopt and find best parameters for models, developed and applied new features
* Trained models on 20 features (out of 50), i used 3-fold KFold, rmse competition metric
* The end result is a mix of lgbm, catboost and knn, find best coef. to mixing model, my result is top 11 percent.

M5 Forecasting-Accuracy(Walmart), Top 8 % | *LGBM, Xgboost, ARIMA, Prophet, Sklearn, Pandas, Numpy*

* Walmart Sales Forecasting
* I invented how validated data (take all data not include 5 moths to validate and predicted for the next month, then this month turns into data and we predict the next, etc. (5 months)
* I make more than 90 features (lags, rolling, time features, etc.), rmse metric, i selected best features use single LGBM model trained on 60 features. One of results is top 8 percent.

Other Competitions: CommonLit Readability(NLP), Coleridge Initiative(NLP) ,Prostate cANcer graDe Assessment (PANDA) Challenge(CV), SIIM-ISIC Melanoma Classification(CV) ...

Education

Murmansk Academy of Economics and Management Murmansk

Economic, economist *Aug. 2002 – May 2008*

Machine learning by Stanford University Coursera, 2018

Machine Learning

Machine Learning and Data Analysis Coursera, 2019

Specialization:

- Search for structure in data, Mathematics and Python for data analysis, Search for structure in data, Applied problems of data analysis

Experience

" I have been into machine learning for over two years now. I was so carried away ML that l decide to change my professional field of activity. I'm looking for a role where i can grow and continue to learn from other experienced team members. "

Technical Skills

Languages: Ruby, Python, SQL Basic knowledge, Russian-native, English-B1
Libraries:LGBM, Xgboost, Catboost, PyTorch, Sklearn, Scipy, Pandas, NumPy, Matplotlib, Seaborn, Plotly
Other: Colaboratory, VScode, Linux, Probability theory, Mathematical statistics, Git