1 Data Motifs Paper

Workloads mentioned in the Data Motifs paper:

Catergory	Application Domain	Workload	Unit of Computation
Deep Learning	Image Recognition Speech Recognition	Convolutional neural network(CNN)	Matrix, Sampling, Transform
		Deep belief network(DBN)	Matrix, Sampling
Graph Mining	Search Engine Community Detection	PageRank	Matrix, Graph, Sort
		BFS, Connected component(CC)	Graph
Dimension Reduction	Image Processing Text Processing	Principal components analysis(PCA)	Matrix
		Latent dirichlet allocation(LDA)	Statistics, Sampling
Recommendation	Association Rules Mining Electronic Commerce	Aporiori	Statistics, Set
		FP-Growth	Graph, Set, Statistics
		Collaborative filtering(CF)	Graph, Matrix
Classification	Image Recognition Speech Recognition Text Recognition	Support vector machine(SVM)	Matrix
		K-nearest neighbors(KNN)	Matrix, Sort, Statistics
		Naive bayes	Statistic
		Random forest	Graph, Statistics
		Decision tree(C4.5/CART/ID3)	Graph, Statistics
Clustering	Data Mining	K-means	Matrix, Sort
Feature Preprocess	Image Processing Signal Processing Text Processing	Image segmentation(GrabCut)	Matrix, Graph
		Scale-invariant feature transform(SIFT)	Matrix, Transform, Sampling, Sort, Statistics
		Image Transform	Matrix, Transform
		Term Frequency-inverse document frequency (TF-IDF)	Statistics
Sequence Tagging	Bioinformatics Language Processing	Hidden Markov Model(HMM)	Matrix
		Conditional random fields(CRF)	Matrix, Sampling
Indexing	Search Engine	Inverted index, Forward index	Statistics, Logic, Set, Sort
Encoding/Decoding	Multimedia Processing Security Cryptography Digital Signature	MPEG-2	Matrix, Transform
		Encryption	Matrix, Logic
		SimHash, MinHash	Set, Logic
		Locality-sensitive hashing(LSH)	Set, Logic
Data Warehouse	Business intelligence	Project, Filter, OrderBy, Union	Set, Sort

2 Motifs Seen in Profiles

- AvgPool - ResNet
- Conv2d - CNN, K-Means, ResNet
- MatMul - KNN
- MaxPool - CNN, K-Means, ResNet
- ReLU - CNN, K-Means, ResNet

3 TensorBoard Profiles

3.1 CNN for Image Recognition on MNIST Dataset

Top 10 TensorFlow operations on GPU

Time (%)	Cumulative time (%)	Category	Operation	TensorCore eligibility	Op is using TensorCore
12.7%	12.7%	Conv2DBackpropFilter	gradient_tape/sequential/conv2d/Conv2D/Conv2DBackpropFilter	✓	✗
10.7%	23.4%	Conv2DBackpropFilter	gradient_tape/sequential/conv2d_1/Conv2D/Conv2DBackpropFilter	✓	✗
9.7%	33%	Conv2DBackpropInput	gradient_tape/sequential/conv2d_1/Conv2D/Conv2DBackpropInput	✓	✗
9.5%	42.6%	Conv2D	sequential/conv2d/Conv2D	✓	✗
6.3%	48.8%	Conv2D	sequential/conv2d_1/Conv2D	✓	✗
4.6%	53.4%	Conv2DBackpropInput	gradient_tape/sequential/conv2d_2/Conv2D/Conv2DBackpropInput	✓	✗
4.3%	57.7%	MaxPoolGrad	gradient_tape/sequential/max_pooling2d/MaxPool/MaxPoolGrad	✗	✗
4.1%	61.8%	MaxPoolGrad	gradient_tape/sequential/max_pooling2d_1/MaxPool/MaxPoolGrad	✗	✗
3.7%	65.5%	ReluGrad	gradient_tape/sequential/conv2d/ReluGrad	✗	✗
3.5%	68.9%	Conv2D	sequential/conv2d_2/Conv2D	✓	✗

3.2 K-Means for Weighted Clustering of MNIST Dataset

Top 10 TensorFlow operations on GPU

Time (%)	Cumulative time (%)	Category	Operation	TensorCore eligibility	Op is using TensorCore
29.1%	29.1%	Unique	Adam/Adam/update_5/Unique	✗	✗
12.8%	41.9%	Conv2DBackpropFilter	gradient_tape/sequential/cluster_conv2d/Conv2D/Conv2DBackpropFilter	✓	✗
7.2%	49.1%	Conv2D	sequential/cluster_conv2d/Conv2D	✓	✗
3.9%	53%	Unique	Adam/Adam/update_2/Unique	✗	✗
3.3%	56.4%	MaxPoolGrad	gradient_tape/sequential/cluster_max_pooling2d/MaxPool/MaxPoolGrad	✗	✗
2.3%	58.7%	ReluGrad	gradient_tape/sequential/cluster_conv2d/ReluGrad	✗	✗
2%	60.7%	BiasAdd	sequential/cluster_conv2d/BiasAdd	✗	✗
1.9%	62.6%	SparseSoftmaxCrossEntropyWithLogits	sparse_categorical_crossentropy/SparseSoftmaxCrossEntropyWithLogits/SparseSoftmaxCrossEntropyWithLogits	✗	✗
1.6%	64.2%	BiasAddGrad	gradient_tape/sequential/cluster_conv2d/BiasAdd/BiasAddGrad	✗	✗
1.4%	65.7%	MaxPool	sequential/cluster_max_pooling2d/MaxPool	✗	✗

3.3 KNN for Image Classification on MNIST Dataset

Top 10 TensorFlow operations on GPU

Time (%)	Cumulative time (%)	Category	Operation	TensorCore eligibility	Op is using TensorCore
15.7%	15.7%	ArgMax	ArgMax	✗	✗
4%	19.7%	MatMul	gradient_tape/sequential/dense_2/Tensordot/MatMul/MatMul	✓	✗
4%	23.7%	MatMul	sequential/dense_1/Tensordot/MatMul	✓	✗
4%	27.6%	MatMul	gradient_tape/sequential/dense_1/Tensordot/MatMul/MatMul	✓	✗
4%	31.6%	MatMul	sequential/dense_2/Tensordot/MatMul	✓	✗
3.9%	35.5%	MatMul	sequential/dense/Tensordot/MatMul	✓	✗
2.4%	37.9%	_Send	IteratorGetNext/_17	✗	✗
1.5%	39.4%	MatMul	gradient_tape/sequential/dense/Tensordot/MatMul/MatMul	✓	✗
1.5%	40.9%	MatMul	gradient_tape/sequential/dense_2/Tensordot/MatMul/MatMul_1	✗	✗
1.5%	42.4%	MatMul	gradient_tape/sequential/dense_1/Tensordot/MatMul/MatMul_1	✗	✗

3.4 ResNet for Image Classification on CIFAR-10 Dataset

	Conv2d	3	3	1545	167	0	0
	BatchNorm2d	3	6	873	313	0	0
	ReLU	3	3	203	70	0	0
	MaxPool2d	3	3	218	83	0	0
+	Sequential	3	0	5287	57	0	0
+	Sequential	3	0	6211	60	0	0
+	Sequential	3	0	5683	32	0	0
+	Sequential	3	0	5708	30	0	0
	AdaptiveAvgPool2d	3	3	510	133	0	0
	Linear	3	3	488	60	0	0

3.5 RNN for Text Classification on IMDB Dataset

Top 10 TensorFlow operations on GPU

Time (%)	Cumulative time (%)	Category	Operation	TensorCore eligibility	Op is using TensorCore
47.5%	47.5%	CudnnRNNV3	cond_40/then/_0/cond/CudnnRNNV3	✓	✗
47.1%	94.5%	CudnnRNNBackpropV3	gradients/cond_grad/If/then/_0/gradients/cond_grad/gradients/cond/CudnnRNNV3_grad/CudnnRNNBackpropV3	✓	✗
0.9%	95.4%	UnsortedSegmentSum	Adam/Adam/update/UnsortedSegmentSum	✗	✗
0.8%	96.2%	ReverseSequence	cond_40/then/_0/cond/ReverseSequence	✗	✗
0.8%	97%	ReverseSequence	gradients/cond_grad/If/then/_0/gradients/cond_grad/gradients/cond/ReverseSequence_grad/ReverseSequence	✗	✗
0.5%	97.5%	Unique	Adam/Adam/update/Unique	✗	✗
0.4%	97.9%	AddN	Adam/gradients/AddN	✗	✗
0.2%	98.1%	ZerosLike	gradients/cond_grad/If/then/_0/gradients/cond_grad/gradients/zeros_like	✗	✗
0.2%	98.3%	ResourceGather	sequential/embedding/embedding_lookup	✗	✗
0.1%	98.4%	_Send	sequential/text_vectorization/RaggedToTensor/RaggedTensorToTensor/_21	✗	✗