DSC 102. Winter 2022. Final Exam.

Instructions:

This is the **graded final exam** in this course. You can take this exam only once.

1. This exam is for **140pts** (+ 10pts extra credit). Your time limit is mentioned in the email.

2. Here is a summary of the questions in order: AI pledge (1pt) + 10 T/F (10 x 2pts) + 15 short conceptual MCQ (15 x 3pts) + 4 short numerical questions (4 x 6pts) + 2 long multi-part-MCQ (2 x 25pts). There is an optional Extra Credit numerical question at the end (10pts).

3. There are **no partial credits** for any (sub-)question in this exam. For every MCQ, pick the single best answer option. For the other questions, **write only your final answer**. Submit scans/photos of only the sheets in this booklet. Use separate sheets for working out the details rough work, etc.--no need to scan/submit those sheets.

Academic Integrity Pledge [1pt]:

I will complete this graded exam in a fair, honest, respectful, responsible, and trustworthy manner. This means that I will complete the exam as if the professor was proctoring my every action. I will act according to the professor's instructions, and I will neither give nor receive any aid or assistance other than what is authorized. I know that the integrity of this exam and this class is up to me, and I pledge to not take any action that would break the trust of my classmates or the professor, or undermine the fairness of this class.

Your full name:

Your PID:

Today's date (in your time zone):

1) True or False Questions [10 x 2pts].

(i) Load balancing helps reduce idle times on processors in a multi-processor setup.

True False

(ii) Shortest Completion Time First is the most fair scheduling policy discussed in class.

True False

(iii) Modin scales much of the Pandas DataFrame API to disk-resident data using a row store approach.

True False

(iv) The main reason to consider locality of reference when executing a data-intensive program is to reduce both cache hits and memory stalls.

True False

(v) Barrier synchronization in bulk synchronous parallelism sometimes causes slowdowns due to straggler workers.

True False

(vi) Task parallelism is the most resource-efficient form of parallelism in terms of amount of DRAM used.

True False

(vii) The logical level of a filesystem works with disk firmware to move bytes to/from disk to DRAM.

True False

(viii) MapReduce/Hadoop ensures fault tolerance through process replication.

True False

(ix) An epoch of Stochastic Gradient Descent is not a commutative aggregate.

True False

(x) AWS S3 is an implementation of shared-memory multi-node parallelism.

True False

2) Pick the single best option for the following MCQ [15 x 3pts].

(i) Which of these factors often reduce(s) the speedups obtained with BSP data parallelism?

A) Overhead of setup and teardown of workers

B) Workers having skewed data partitions

C) Communication between manager and workers

Both A and B

All of A, B, C

None of the rest

(ii) Which of the following is/are major cons of using NFS (vs. HDFS)?

A) Very difficult to setup and use

B) Not scalable to very large files

C) Contention for concurrent reads and writes

Both A and B

Both A and C

Both B and C

(iii) Which of the following is not true for scalable SGD?

Works well for non-convex losses too

Need to randomly shuffle data

Does a filescan per epoch

Shuffling data once upfront often suffices

Samples mini batch from data with replacement

None of the rest

(iv) Which of these scheduling policies often causes short jobs to wait unfairly long times for their turn?

Last in First Out

Shortest Job First

Round Robin

Random

Shortest Completion Time First

First in First Out

(v) Which parallel data system is based on the shared-nothing parallelism paradigm?

Greenplum

Redshift

Hadoop

Spark

All of the 4 listed

None of the 4 listed

(vi) What is the technical term for preemptively reading bytes from DRAM addresses not explicitly asked by a program?

Caching

Cache miss

Prefetching

Cache hit

Register spill

Page fault

(vii) Which of the following feature engineering techniques do(es) not preserve the semantics of the features?

A) Feature subset selection

B) PCA

C) Matrix factorization

Both A and B

Both B and C

Both A and C

(viii) Which of the following devices in a typical memory/storage hierarchy is the costliest in terms of per-byte price?

Magnetic disk

CPU registers

Cache

Solid state drive

Hard disk drive

DRAM

(ix) Which of the following is true about Parameter Server?

Each worker handles a part of the model parameters at a time

Model params may get out-of-sync or stale

Network I/O cost is high

Workers typically send gradients to manager for updates at each mini-batch

None of the 4

All of the 4

(x) Which structured data model(s) support(s) the transpose operation?

A) Relation

B) Matrix

C) DataFrame

Both A and B

Both B and C

Both A and C

(xi) How many bits are needed to distinguish 784 data items?

7

8

9

10

11

None of the rest

(xii) What is the turnaround time for a process in the scheduling context?

Start time - Arrival Time

Completion time - Response time

Completion time - Arrival time

Completion time - Start time

Response time - Arrival time

None of the rest

(xiii) Which of these is considered IaaS in cloud jargon?

Aurora

SageMaker

Redshift

EC2

Chime

DynamoDB

(xiv) What is the term for a virtual slice of DRAM assigned to a given process?

Process space

Address space

Page space

State space

Virtual space

None of the rest

(xv) Which of the following tools is/are specifically built for running tensor dataflow graphs?

A) Spark

B) PyTorch

C) Hadoop

Both A and B

Both A and C

Both B and C

3) (6pts) Floating point formats

Suppose you magically have a base-3 computer. Consider this custom float6 representation using base-3 digits.

	sign	exponent			fraction	
	x	x	x	x	x	x
Digit index:	5	4	3	2	1	0

$$(-1)^{sign} \times 3^{exponent-4} \times \left(1 + \sum_{i=1}^{3} d_{3-i} 3^{-i}\right)$$

What is the *second largest* number representable in this float format? Ignore special cases. Enter its real-valued decimal representation. Hint: It is an integer.

FINAL ANSWER:

4) (6pts) Multi-core parallelism

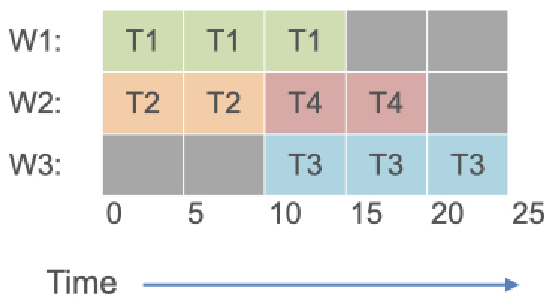
Suppose you parallelize a data-intensive program across multiple cores on a single machine. 40% of the program is inherently sequential, while the rest can yield perfect linear speedup for any number of cores.

The single-core runtime of the program is 60min. What is the runtime (in minutes) of parallel 4-core execution? Write only the exact decimal number without preceding or trailing zeroes.

FINAL ANSWER:

5) (6pts) Task-parallel scheduling

Consider this workload with 4 tasks being run in a task-parallel manner on 3 workers. Tasks T3 and T4 depend on T2.



What is the lowest possible task-parallel completion time of this workload given only 2 workers?

FINAL ANSWER:

6) (6pts) Model selection costs

Suppose you try an MLP on a tabular dataset with 20 features. As part of model selection you decide to try all subsets of size 19 features (apart from the full feature set) and 2 different choices of neural architecture.

For hyperparameter tuning you use random search with 25% of the combinations in the grid created with the following hyperparameter settings: 4 different learning rates and 4 different L2 regularizers. You train all models for 10 epochs each.

What is the total number of passes over the training dataset in this workload?

FINAL ANSWER:

7) MapReduce for Feature Ranking

You are given a large training dataset with schema *D (Y, X1, X2, …, Xd)*. The target *Y* and each feature are given to be boolean, each represented as a 0 or 1 with a 1-byte short integer data type. The number of examples is *n*. The number of features *d* is only in the few hundreds.

The dataset is stored in row-major format and is uniformly sharded row-wise (no compression) on an HDFS cluster that has *k* worker machines and does not use any replication.

Devise a single MapReduce job to rank all features based on their *mutual information* with the target. It must be scalable along the number of examples. Given feature *X* and target *Y*, their mutual information, denoted *I (X; Y)*, is mathematically defined as follows as explained here: https://en.wikipedia.org/wiki/Mutual_information

$$I(X;Y) = \sum_{x\in\{0,1\}} \sum_{y\in\{0,1\}} P(X=x,Y=y) log\frac{P(X=x,Y=y)}{P(X=x)P(Y=y)}$$

(i) (5pts) Given (n, d, k) = (5 billion, 100, 10) what is the storage footprint of the dataset on each worker in GB?

25

50

100

250

500

1000

None of the rest

(ii) (5pts) What is the most accurate time complexity of the Map stage on a whole shard in terms of n, d, and k?

O(n)

O(dn)

O(dkn)

O(dn/k)

O(d/(kn))

O(kn/d)

None of the rest

(iii) (5pts) What is the expression for the number of sufficient statistics emitted by each Map process in terms of n, d, k? Assume that all relevant counts for all boolean values of a variable are calculated separately.

2d+1

2d+3

3d+1

3d+3

4d+1

4d+3

6d+1

6d+3

None of the rest

(iv) (5pts) What is the most accurate time complexity of the Reduce stage in terms of n, d, and k?

O(d)

O(k)

O(dk)

O(dn/k)

O(n/k)

O(d/k)

None of the rest

(v) (5pts) What is the most accurate space complexity of the state information emitted by the Reduce process in terms of n, d, k?

O(1)

O(n)

O(d)

O(k)

O(dn)

O(dn/k)

None of the rest

8) Scalable Feature Engineering and Model Selection

In this "grand finale" question, you will apply concepts from across topics to figure out the costs of scaling feature interactions for amplifying the accuracy of a linear model with a multi-step workflow.

You are given a large training dataset with schema *D (Y, X1, X2, …, Xd)*. The target *Y* and all features are numeric and stored as float64.

The number of examples is *n*. The dataset is laid out in row store format on disk (no compression) on your single machine. No batching of filescans or caching of data is performed.

Step 1) You build linear models using BGD. You try 3 different learning rates and train all models for 5 epochs each.

Step 2) You now read the data and append quadratic (order 2) feature interactions as explained in the lecture. You write out the output file to disk. All features are still stored as float64.

Step 3) You rebuild the linear models on the output file of step 2 using the same model selection process as the first step.

Given: (n = 1 billion, d = 3)

(i) (4pts) What is the size in GB of the file D on disk? (Hint: Do not forget to count Y.)

FINAL ANSWER:

(ii) (5pts) What is the total I/O cost in GB of step 1 for reading data files?

FINAL ANSWER:

(iii) (6pts) What is the size in GB of the file written out in step 2? (Hint: Do not forget to count Y.)

FINAL ANSWER:

(iv) (5pts) What is the total I/O cost in GB of step 3 for reading data files?

FINAL ANSWER:

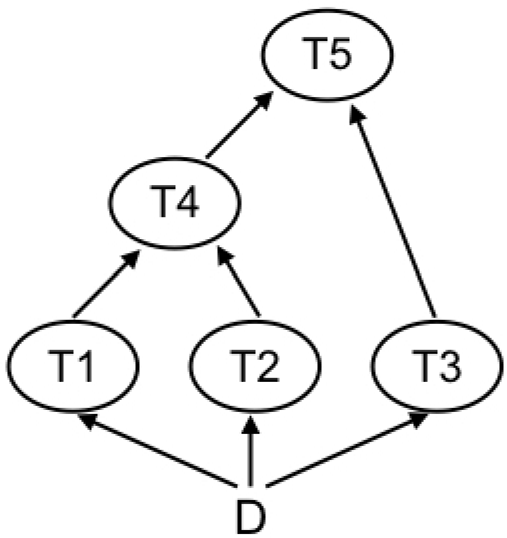
(v) (5pts) What is the degree of parallelism of this entire 3-step workflow if you were to run it in a task-parallel manner?

FINAL ANSWER:

(10pts) Extra credit question:

You are given the following task graph. The runtimes of the tasks are given below. Each task's work is amenable to perfect linear speedup on multiple workers with BSP data-parallelism, albeit with 2 units of manager overhead each before and after the actual parallel worker execution of the task's work.

You are given a homogeneous cluster with n machines. What is the lowest value of n at which a fully BSP-parallel execution of this workload becomes strictly faster than a fully task-parallel execution on that sized cluster?



Runtimes of (T1, T2, T3, T4, T5) = (10, 10, 15, 15, 5)

FINAL ANSWER: