BACC 2022

Team Micronsoft

Cao Yuxuan, Jin Cheng Hao, Lee Jing Hong, Lim En Hao

Hwa Chong Institution

Question 1

Micron's role & Impacts on Micron

Micron's Role

Redundancy in operations

- **Globally diversified production has been very helpful because it mitigates the COVID-19 protocols that were impacting production in certain areas**

Micron's Role

"US 150 billion, 10-year investment plan"

- **US Fab expansions and working with governments around the world**
- **Have more autonomy over the sourcing of raw materials and distribution of products.**
- **Help to alleviate shortage by injecting more supply**

Micron’s Role

Pivotal Role in the retail sector

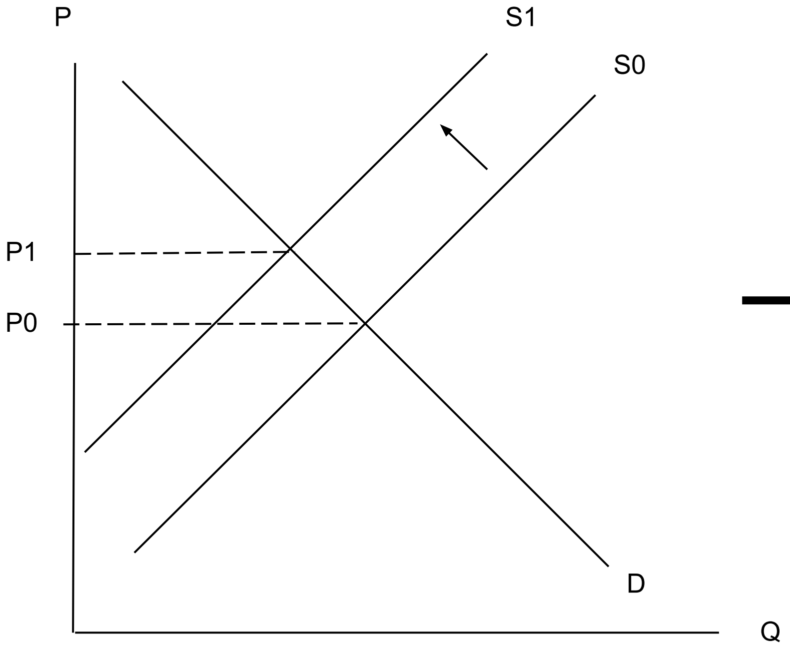
- Many of its major customers depend heavily on the performance of their retail stores
- Apple suffered delays in production due to the shortage
- Production of its latest model was slashed

Impact on Micron

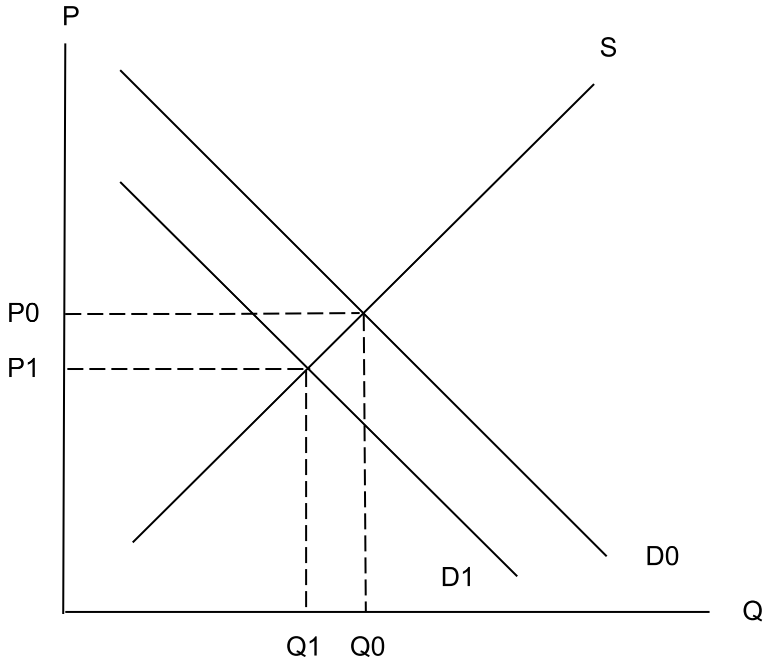
Disrupted supply chains for complementary goods

- **Micron's DRAM and NAND chips need to be paired with computing-purpose chips.**
- **This has affected demand for Micron's products in general.**
- **Tesla has also recently removed USB ports due to the shortage.**

Impact on Micron

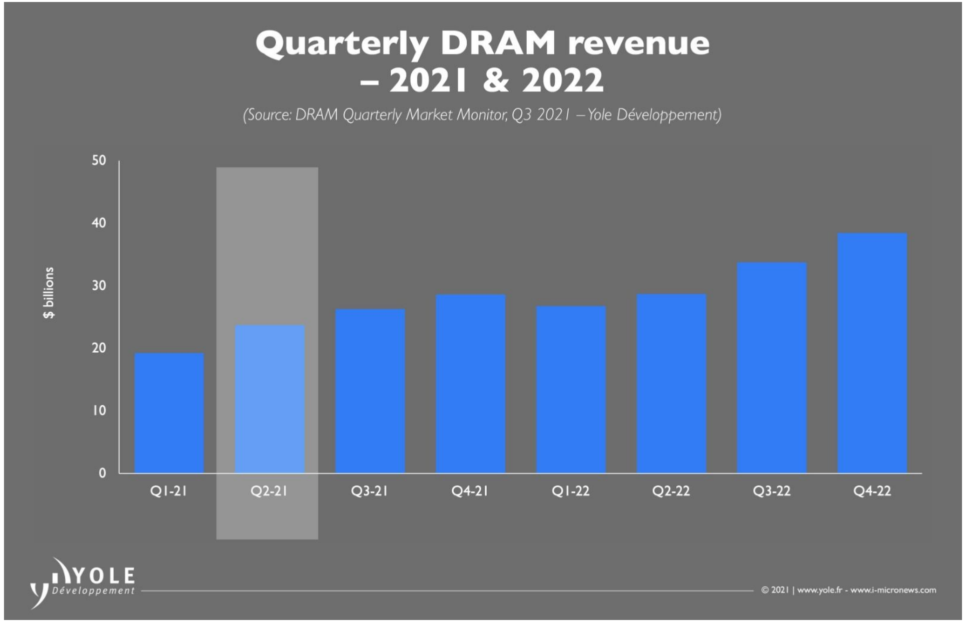


Market for complementary chips



Market for DRAM and NAND, Revenue falls

Impact on Micron



Impact on Micron

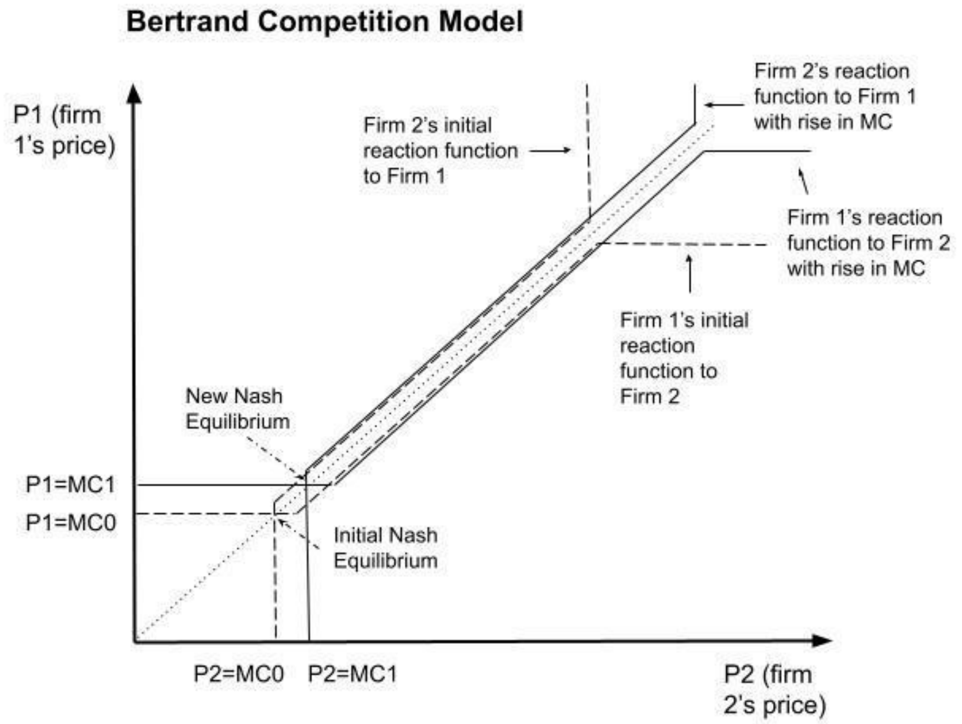
Expanding production capacity

- Micron does not have a fabless model
- Not dependent on foundries like TSMC
- Avoid large premiums to reserve production
- Able to expand on production capacity

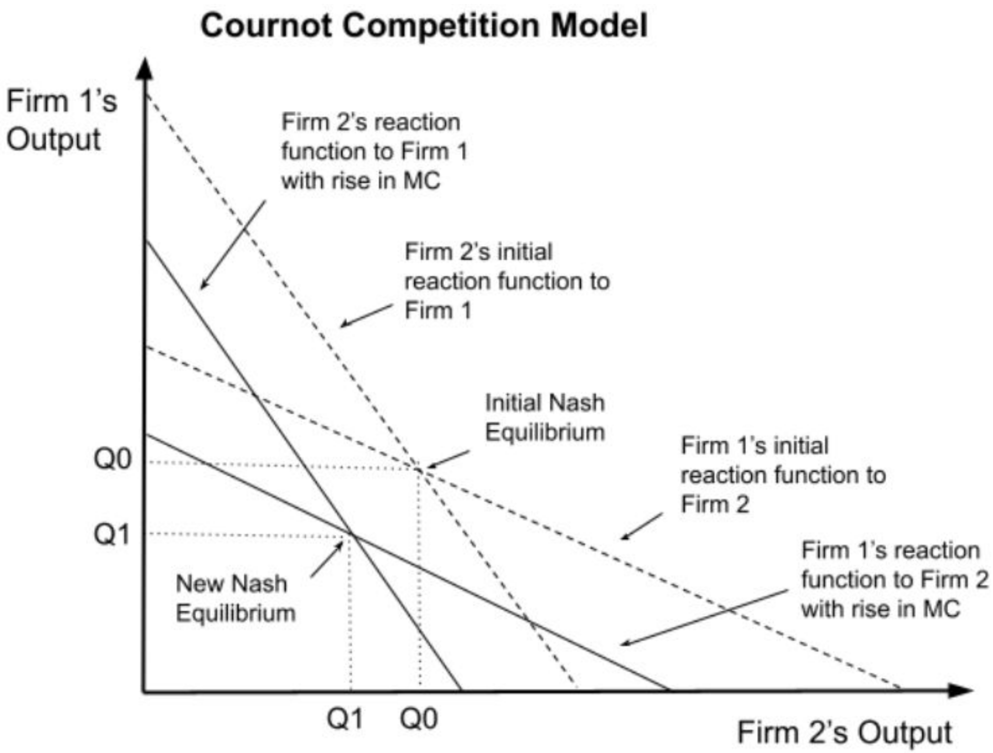
Impact on Micron

- However, the cost of error for Micron and similar chip production companies has skyrocketed during this shortage, with disruptions incurring greater operational costs due to the need for better emergency planning and response.
- Incident of contaminated flash chips from Western Digital factory. WD was forced to respond to sudden cost increases by raising the prices of its affected products.

Impact on Micron



Impact on Micron

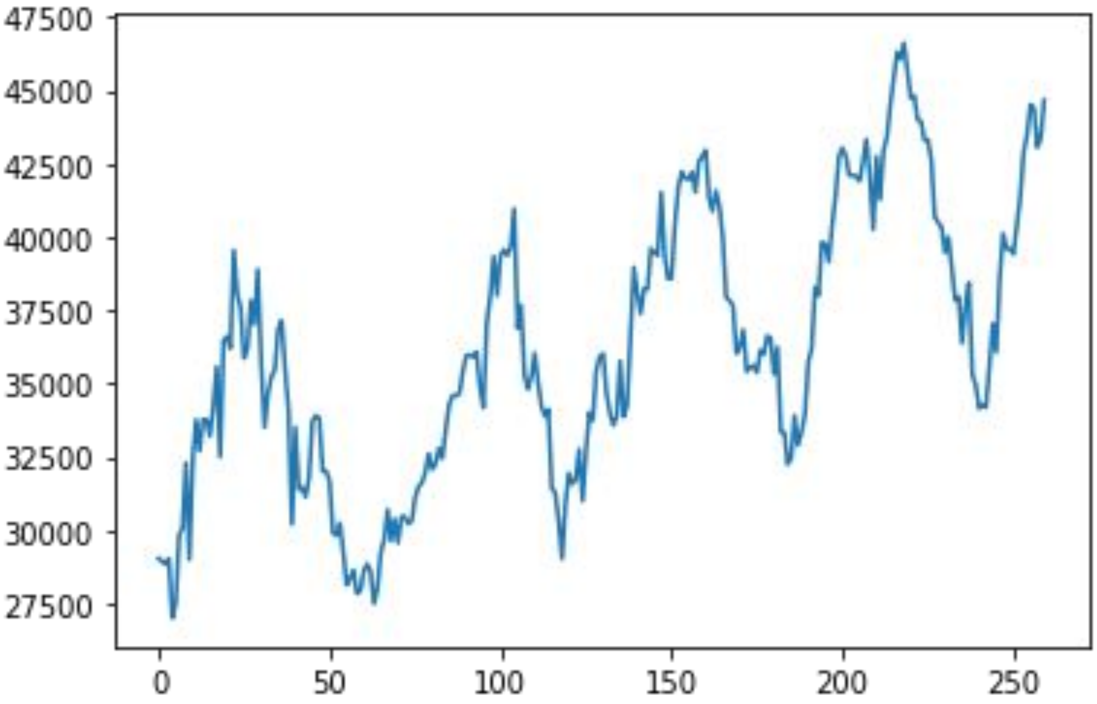


Question 2.1

Demand Forecasting

Overview of data

- Seasonal trend
- Upward trend
- Noise



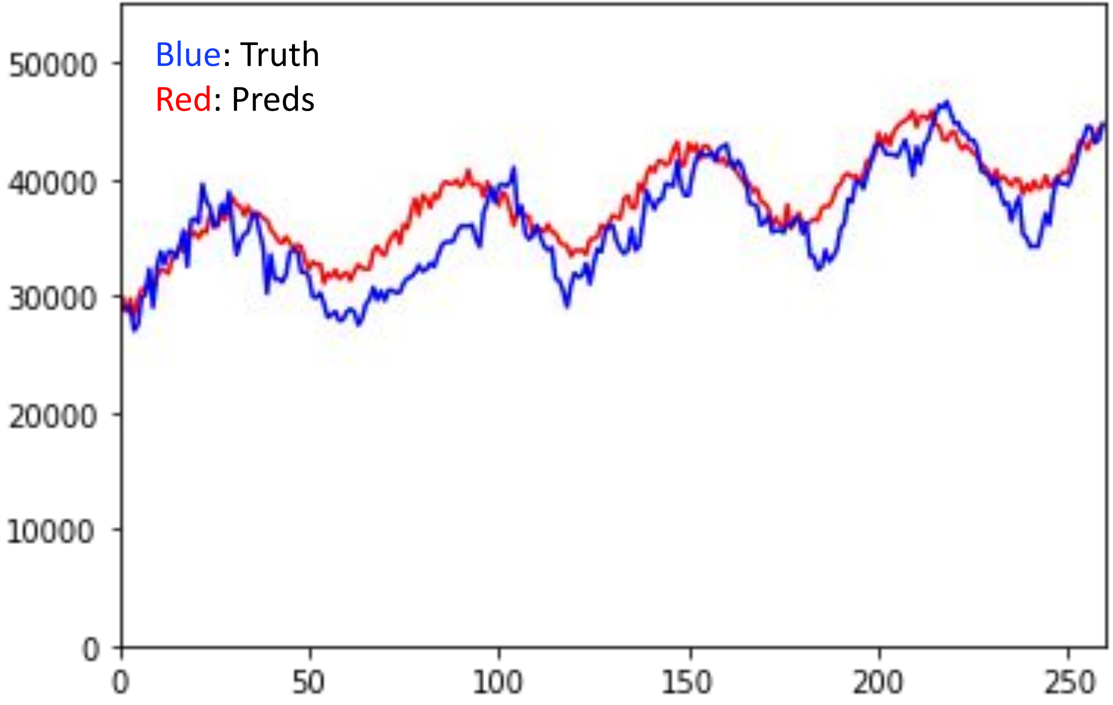
Mathematical Approach

$$P(x) = ax - k\left[cos(\frac{2\pi}{p}x) - cos(\frac{2\pi}{p})\right] + E$$

- ***P(x)*** := prediction function, where the argument is the # of the week being predicted
- ***a*** := gradient of slope = change in deseasonalized demand data over ***N*** timesteps / ***N***
- ***k*** := amplitude of curve = 0.5 * average peak to trough distance
- ***p*** := period of curve = 60 (by observation)
- ***E*** = random variable that follows a Gaussian distribution defined by the standard deviation of error terms between the given values and their 15-week moving average

Mathematical Approach

- MAE: ~**2200**
- Troughs are sharper while peaks are rounder
- Will become increasingly unreliable

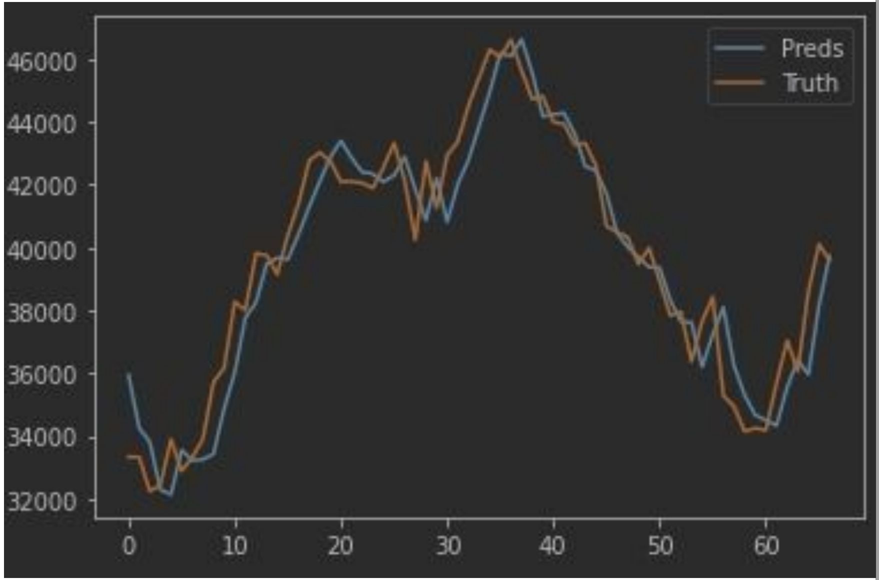


Deep Learning Approach

- Universal Approximation Theorem
- Large information receptive field
- High degree of non-linearity
- Gradient Descent
- Can be retrained in future and reused on other products
- Performance scales with amount of data
- RNN (Recurrent Neural Network) - failed to perform well
- MLP (Multi-layer Perceptron)

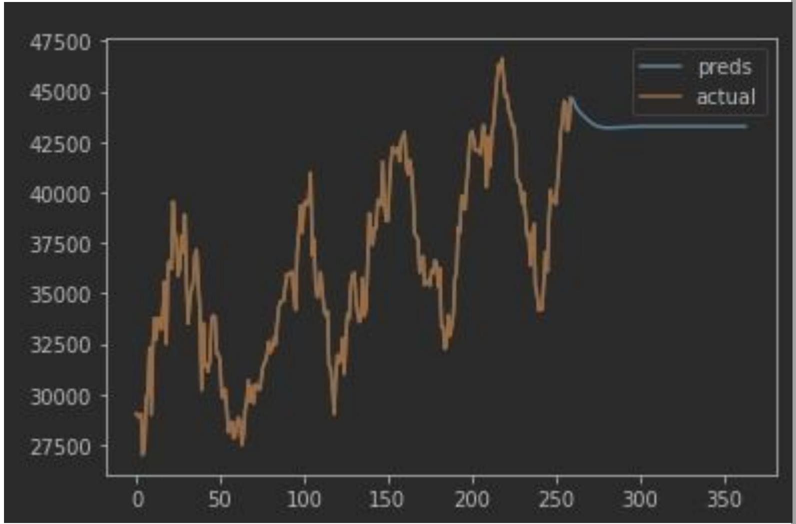
Deep Learning Approach - RNN (LSTM)

- Input: past 52 weeks of data in **sequence**
- Output: demand predicted for next week
- Split data into train and validation
    - 4 years for train, 1 year for validatation
- MAE: **970** on validation set
- Model learnt false info of copying previous value
    - Due to dataset and loss function



Deep Learning Approach - RNN (LSTM)

- Did not work when put to test
- There must not be simple sequential pattern
- (a, b, c, d) cannot have a relationship where a+b=c, b+c=d etc.

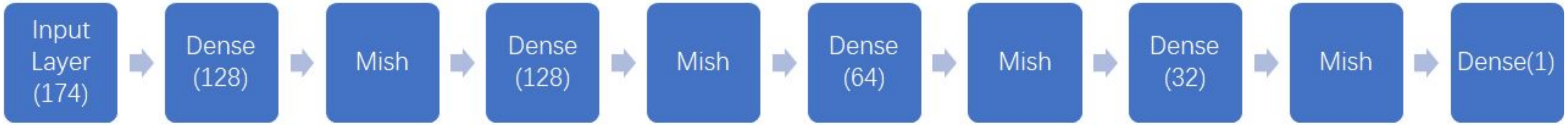


Deep Learning Approach - MLP

- Input sequence is **not** considered thus model will not be naturally biased to more recent data
- Prevents ‘copying’ issue in RNN
- Consists of Fully Connected (FC) layers
- Activation functions in middle to enhance non-linearity

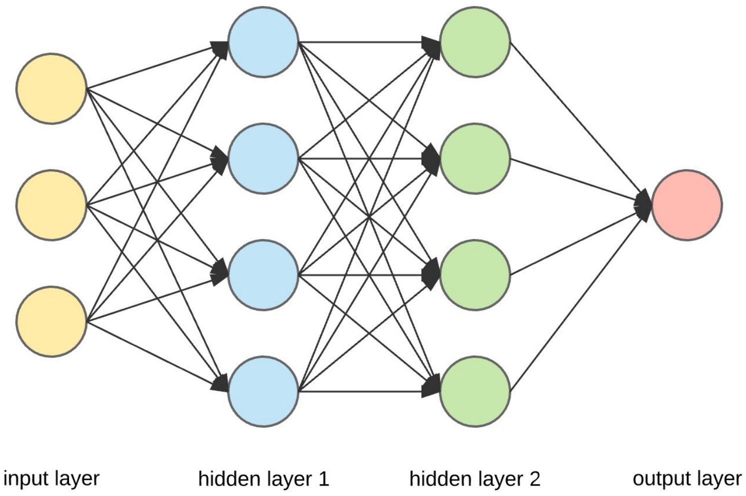
Deep Learning Approach - MLP

- Input: past 3 * N weeks where N = period = (52, 58)
- Output: demand predicted for next week
- Consists of fully-connected (FC) layers and Mish activation in between them
- MAE on full data: **244** (N=58), **269** (N=52)
- Model ensembling by taking mean of both model's predictions
  - Reduces noise

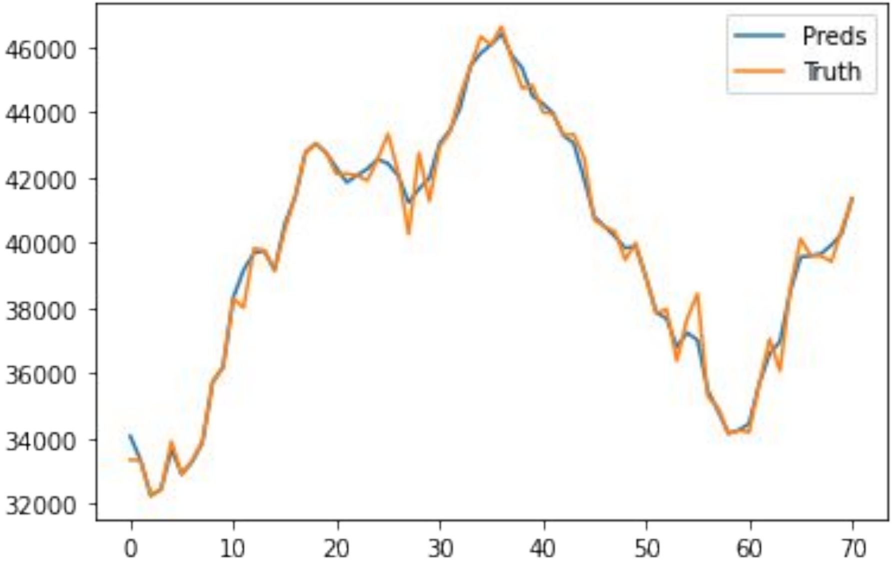


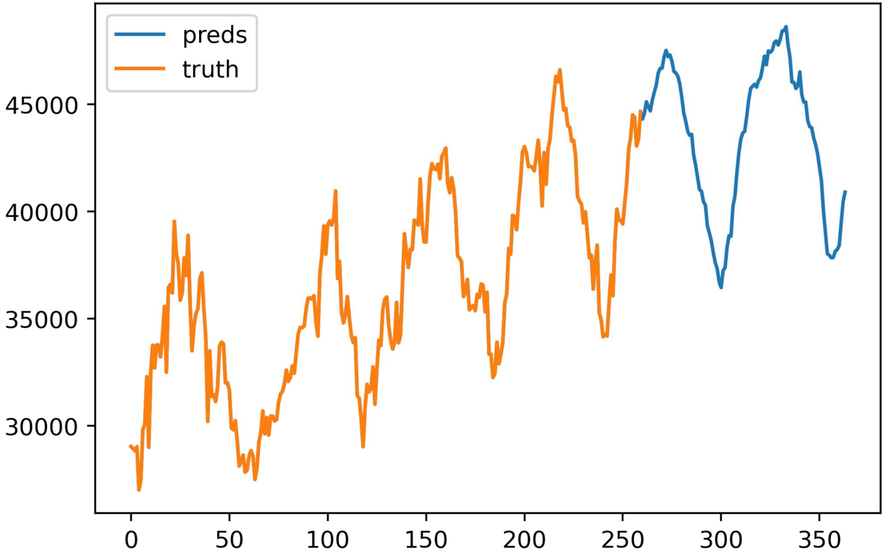
Fully Connected Layers (Dense)

- y = w · x + b
- A single layer contains many neurons
- Output fed to each and every neuron in the next hidden layer
- Activation function is applied to all outputs before feeding into next layer
- Mish activation: $mish(x) = x\ tanh(ln(1 + e^{x}))$

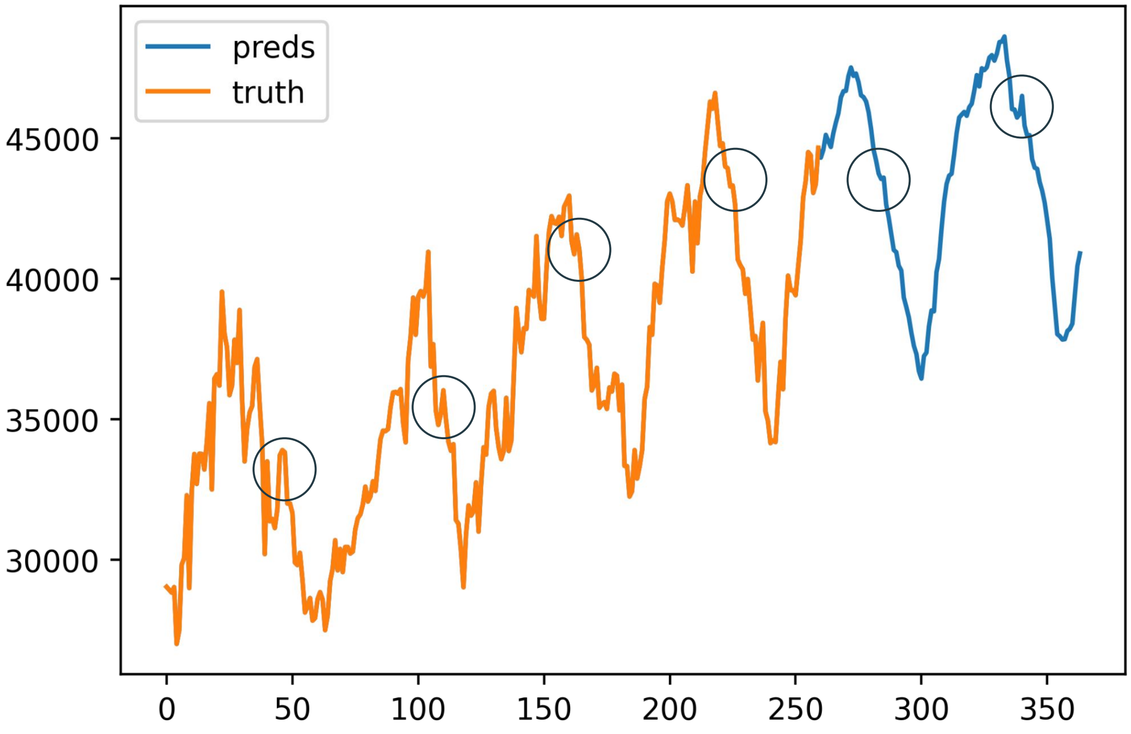


MLP Performance

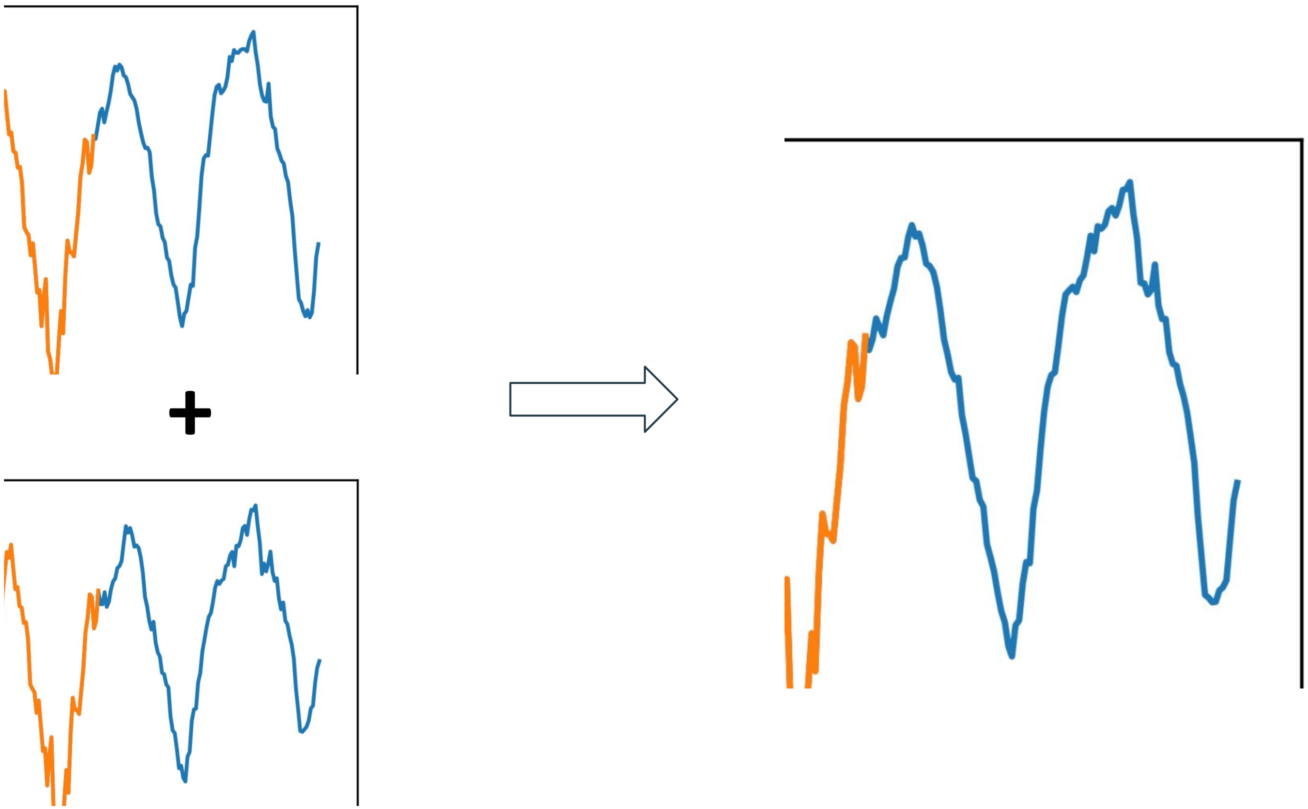




MLP Performance



MLP Ensembling



MLP Limitations

- ‘Black box’: lack of explainability for human understanding
  - Provide little value for human analysis on results
  - Hard to evaluate effectiveness into the future
- Random nature of stochastic gradient descent
  - Not 100% reproducible
  - Precautionary measures possible
- Fixed input size
  - Reusability depends on amount of data available
  - Adjustments to model architecture may be needed
- Sensitive to data quantity
  - Scales with # of input features (historical data inputted at once)
  - May underperform mathematics model when data is too limited
  - Negligible for this particular task

Question 2.2

Minimum Machine Requirement

Formula for Machine Capacity

Combining all the given equations, we get an expression for the Machine Capacity of an individual machine:

Machine Capacity (*M*)= Availability (*A*) * RPT Basis (R_B) * [Minutes in a week * Utilization (*U*) / Total RPT (R_T)] * Load Size (*L*) = $10080AR_BUL/R_T$

We know that ***A, U, L,*** and R_T are **constants** for each type of machine. Therefore, ***M*** is a function of R_B.

Simplifying Assumption

Premise 1: Machines are capable of performing their individual tasks irregardless of the progress of other machines.

Premise 2: Machines can only take one task at a time.

Assumption: Machines must finish all their given tasks in each iteration of the manufacturing process. Thus, we can let $\boldsymbol{R_B}$ equal the **total RPT basis across all steps for each machine type** for each iteration of production and then calculate $\boldsymbol{M}$.

Formula for # of Machines Required

Implication: The number of machines of a certain type required to fulfil the demand is:

Ceil(Forecasted demand in Week N / *M*)

Example: (Cells are highlighted in green if the current number of machines can supply the forecasted demand and red if there will be a bottleneck in this type of machine)

		Albatross	Cattle	Dragon	Flamingo	Grouse	Herring	Peacock	Plovers
Week	**Demand**	**Number of Machines Required (for each Machine Type) to Meet Forecasted Weekly Demand**							
1	44315	187	276	102	259	83	118	455	575

Question 2.3

Optimizing for Meeting Demand

27	42648	180	266	98	249	80	113	438	554
28	42178	178	263	97	247	79	112	434	548
29	41607	176	259	96	243	78	110	428	540
30	41029	174	256	94	240	77	109	422	533
31	40951	173	255	94	240	77	109	421	532
32	40467	171	252	93	237	76	107	416	526
33	40297	170	251	93	236	76	107	414	523
34	39332	166	245	90	230	74	104	404	511
35	38995	165	243	90	228	73	104	401	506
36	38626	163	241	89	226	73	103	397	502
37	38084	161	237	88	223	72	101	391	495
38	37616	159	234	86	220	71	100	387	489
39	37327	158	233	86	218	70	99	384	485
40	36716	155	229	84	215	69	98	377	477
41	36448	154	227	84	213	69	97	375	473
42	37249	158	232	86	218	70	99	383	484
43	37375	158	233	86	219	70	99	384	485
44	38318	162	239	88	224	72	102	394	498
45	38872	164	242	89	227	73	103	400	505
46	38843	164	242	89	227	73	103	399	504
47	40238	170	251	92	235	76	107	414	523
48	40710	172	254	94	238	77	108	418	529
49	41789	177	260	96	244	79	111	430	543

Assumptions

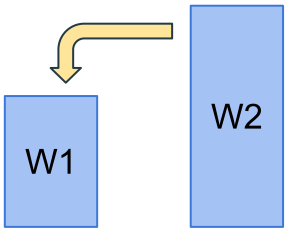
- There is no difference if Micron chooses to buy machines over different periods of time or if they buy the total number of extra machines required in the first week
  - This may not be reflected in the real world as companies have cash flow constraints and risks to consider
- Micron wants to spend as little money as they can on buying machines
- The demand for chip **types** within the 2 year period will not change **drastically**, which would thus allow Micron to produce chips in advance based on predicted demand

Naive Solution

Albatross	Cattle	Dragon	Flamingo	Grouse	Herring	Peacock	Plovers
15	21	19	63	16	0	0	36

- Micron buys all their required machines in the first week
- Take data from 2.2
- Maximum # of machines required starting from Week 38 - original #
- Reasoning:
    - Up to Week 38, Micron **will not be able to meet demand** due to a bottleneck in Flamingo
    - After Week 38, there will be a second spike in demand starting from approximately Week 44
    - The total buying and installation time for every machine type will come before their respective machine requirement spikes.

Advanced Solution



Naive:

Albatross	Cattle	Dragon	Flamingo	Grouse	Herring	Peacock	Plovers
15	21	19	63	16	0	0	36

With production-shifting:

Albatross	Cattle	Dragon	Flamingo	Grouse	Herring	Peacock	Plovers
0	0	8	41	10	0	0	0

Concept: Production shifting

Under the assumptions, we assume that Micron can pre-produce chips should they predict sufficient demand in the future. Thus, they can shift some production from **later** weeks to **earlier** weeks in order to lower the number of machines required in the later weeks. This raises productivity and decreases costs for Micron.

Optimal Production Shifting

Let (X, Y) be the # of machines demanded in (week S, week S+1), where X > Y.

Assume $\exists\, M \in \mathbb{R}^+$ such that it represents the optimal distribution over a certain subset of weeks where we can theoretically use a fractional number of machines. When we reach week (S+1), assume that it is **still in the given subset of weeks**; therefore:

$|M - X| < 1$ & $|M - Y| < 1$.

Therefore, if we shift any positive integer amount of production from Y to X, and we assume that $|M - X| < 1$, then it must be the case that $M - X + 1 < 1$, and thus:

$M - Y - 1 < -1 < M - Y$. This is a contradiction as we assumed $|M - Y - 1| < 1$, therefore week (S+1) is not in the given subset of weeks where M is the optimal distribution.

THANK YOU