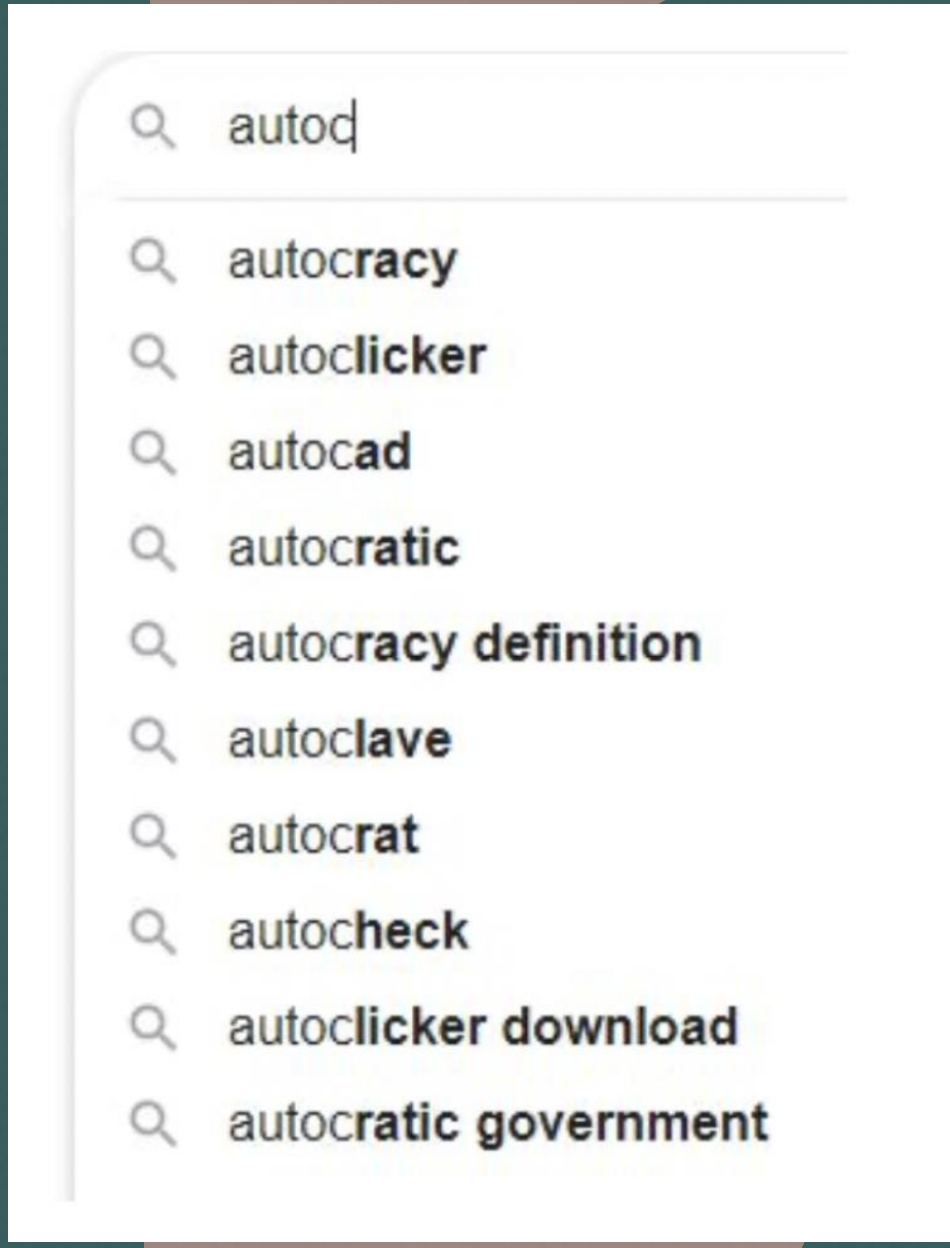
Chatbot Recommender

Max Everest, Nirjhar Deb, Nicholas Jaimes,
Rish Desai, and Christina Chen

Autocomplete Idea

- Give students recommendations on what to ask.
- Think Google search autocomplete for students.
- Last semester, we worked on building a dataset.
- This semester, we worked on creating algorithms.



Real Piazza Questions

- Pulled questions from Piazza for our dataset.
- Played with 180 questions from Fall 2020
- Can be missing some context (ex. How do I solve 3.2?)
- Structure:


```
{
    "subject":"Sinusoids Multiplication Formula Order",
    "content":"When going from multiplication to addition of sinusoids...
    "created":"2020-09-18T14:45:48.732Z",
    "views":11
},
```

AI Generated Mock Questions

- Generated mock questions using AI.
- For some algorithms, we will use these in place of real data.
- This way, we are not limited by the size of the dataset.
- Structure:

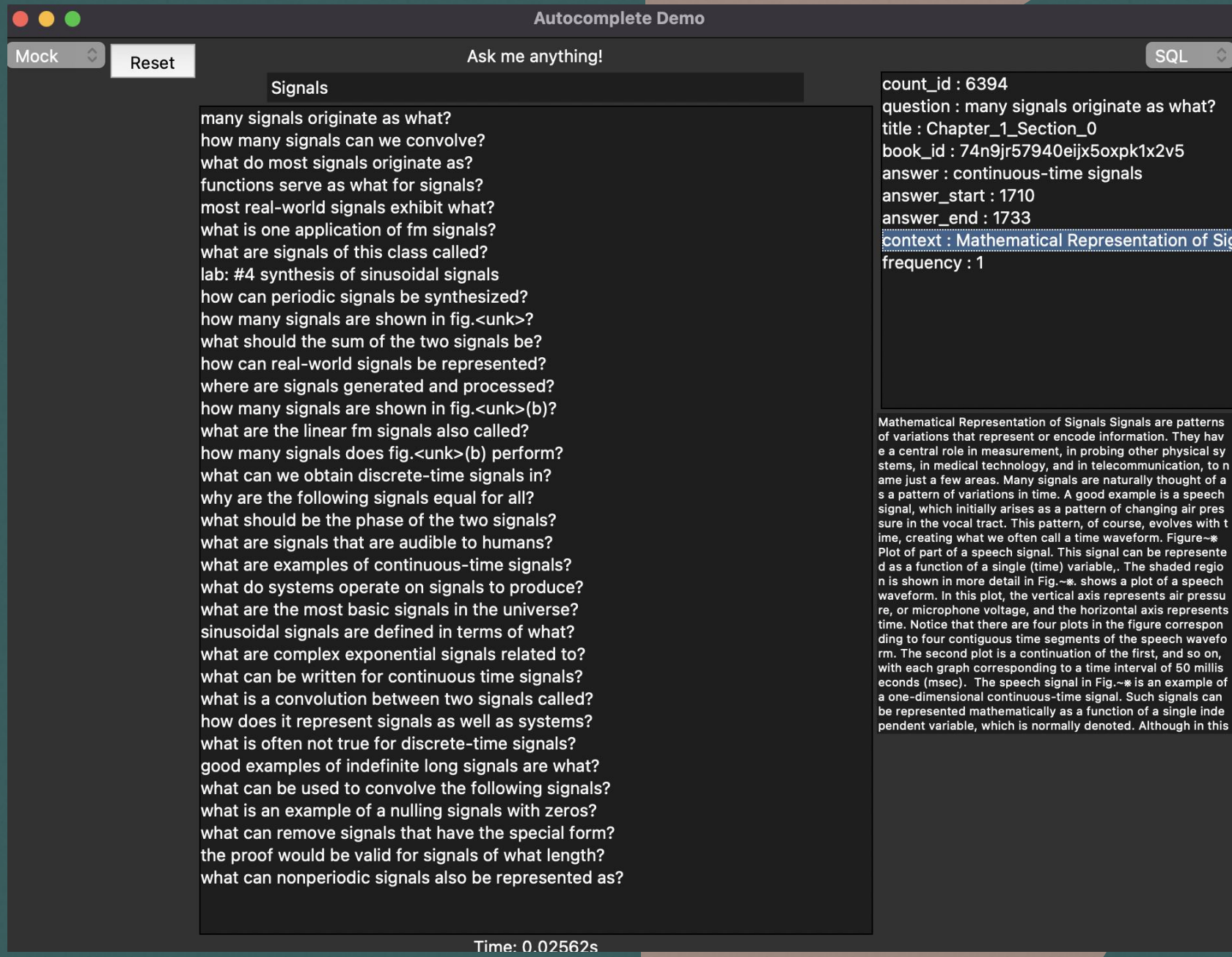
```
{
    "question": "what is the minimum sampling rate of a continuous-time signal called?",
    "chapter": "Chapter_4_Section_0",
    "frequency": 3,
    "time_stamp": 42,
    "answer_start": 1103,
    "answer_end": 1115
},
```

Intended Audience

- Each algorithm is meant to provide a better experience for students and help them find more efficient ways to learn the course material
- Through each algorithm there's a new purpose for reaching this goal of efficiency

GUI

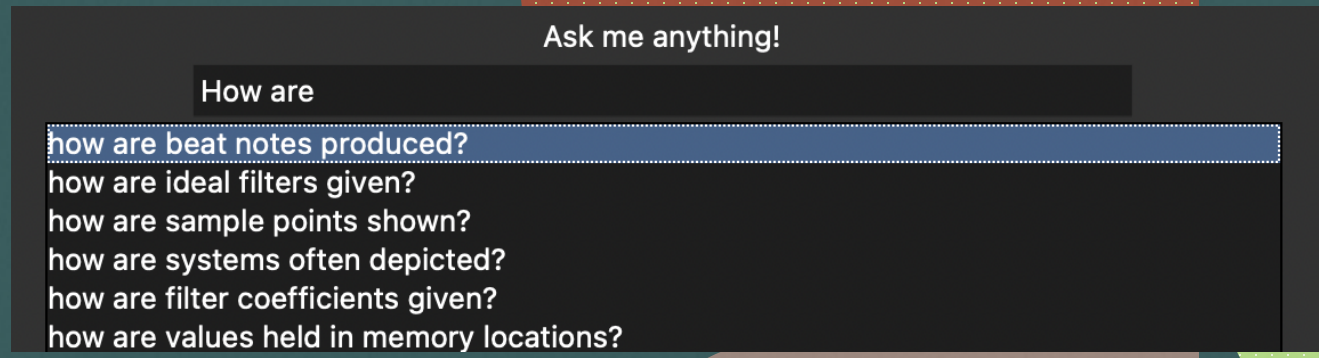
- Interactive GUI that can visualize each algorithm in real time!
- Supports:
  - Algorithm Selection
  - Dataset Selection
  - Context Display
  - Inspection Window (for blobs of text)
  - Algorithm Timer



Recommender Algorithms

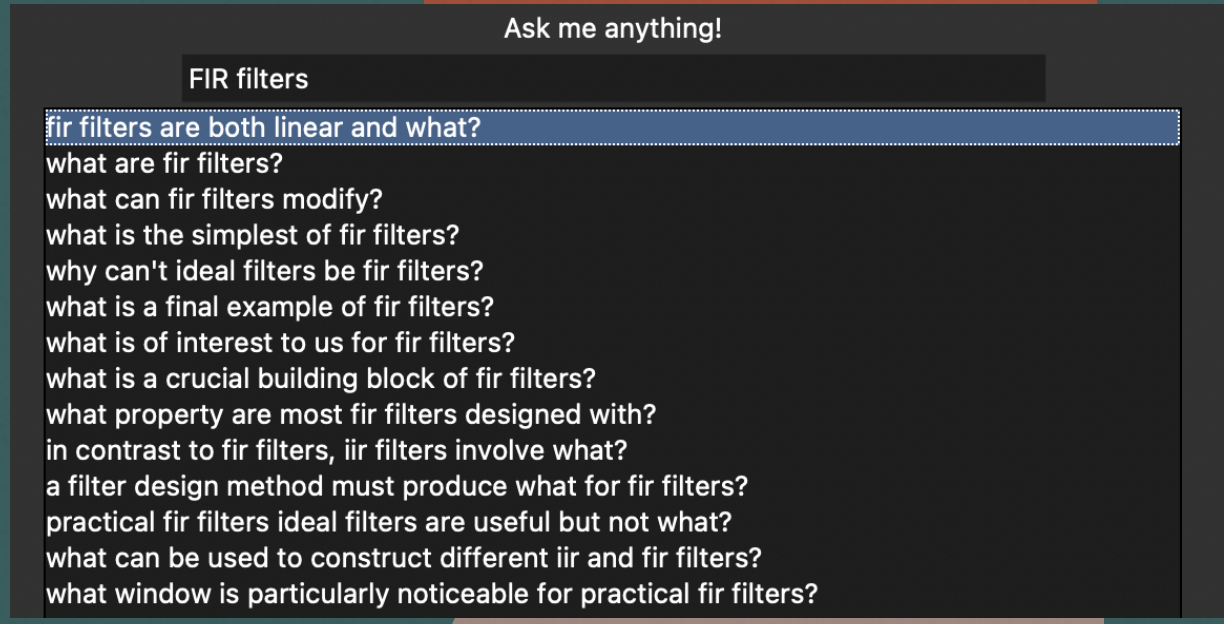
SQL Prefix Matching

- Searches the entire dataset for matching prefixes.
- Input – What d
- Suggestions:
  - What does rad/s mean?
  - What does FIR stand for?
  - What dictates Newtons Second Law?
- Extremely fast, but narrow results.



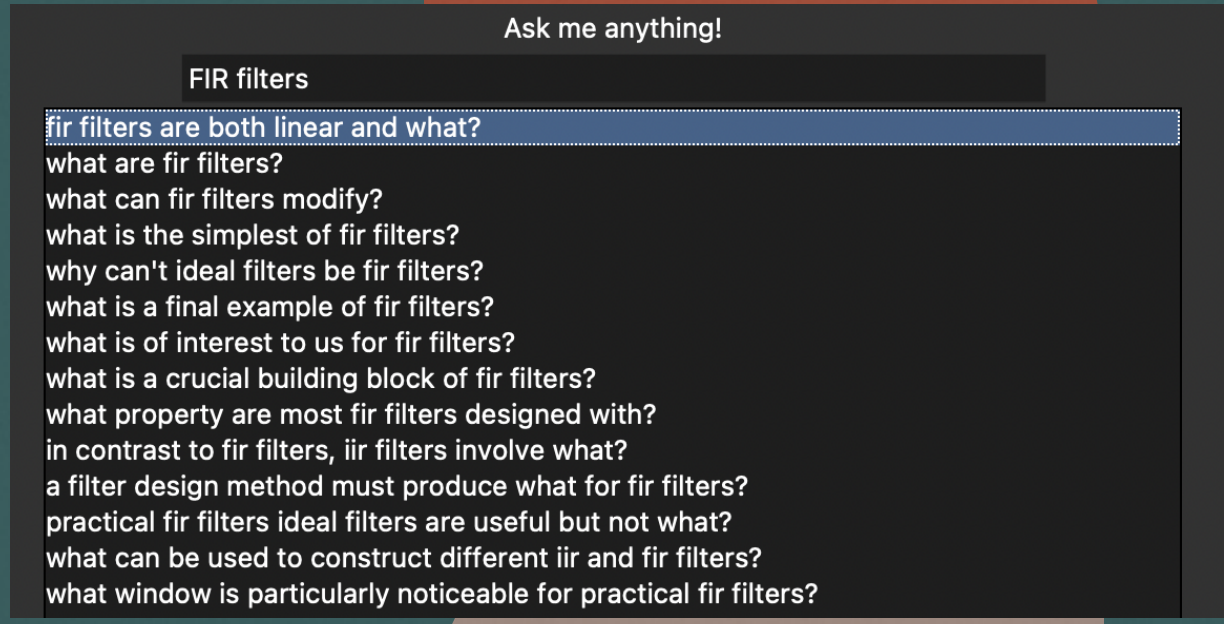
SQL Pattern Matching

- Searches the entire dataset for any pattern matches.
  - Boyor-Moore
- Input – Signal
- Suggestions:
  - Signals tend to originate as what?
  - What is a continuous signal?
  - What is a discontinuous signal?
- Wider, although still narrow range of results.



SQL Algorithm

- Combines prefix matching and pattern matching into one algorithm
- Prioritizes direct prefix matches
- Once prefix matches run out, will output general pattern matches.
- Sort Order:
  - Match Category (prefix then pattern)
  - Ascending Frequency
  - Descending Length

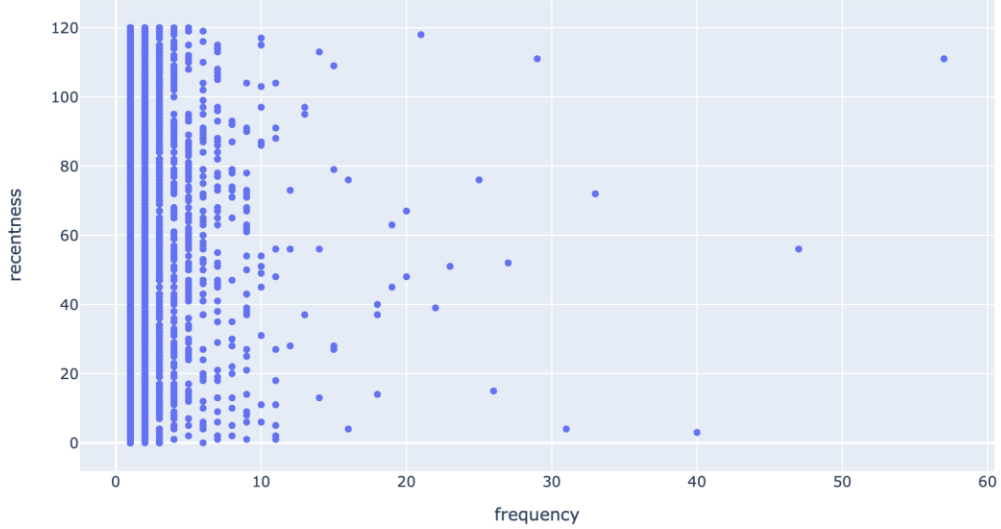


Mork Data Features

- Generated from the Stanford Dataset last semester
- Includes near 6000 questions
- Created frequency and time_stamp features
- Frequency stands for number of views
- Time_stamp stands for the day counted from the first day of class

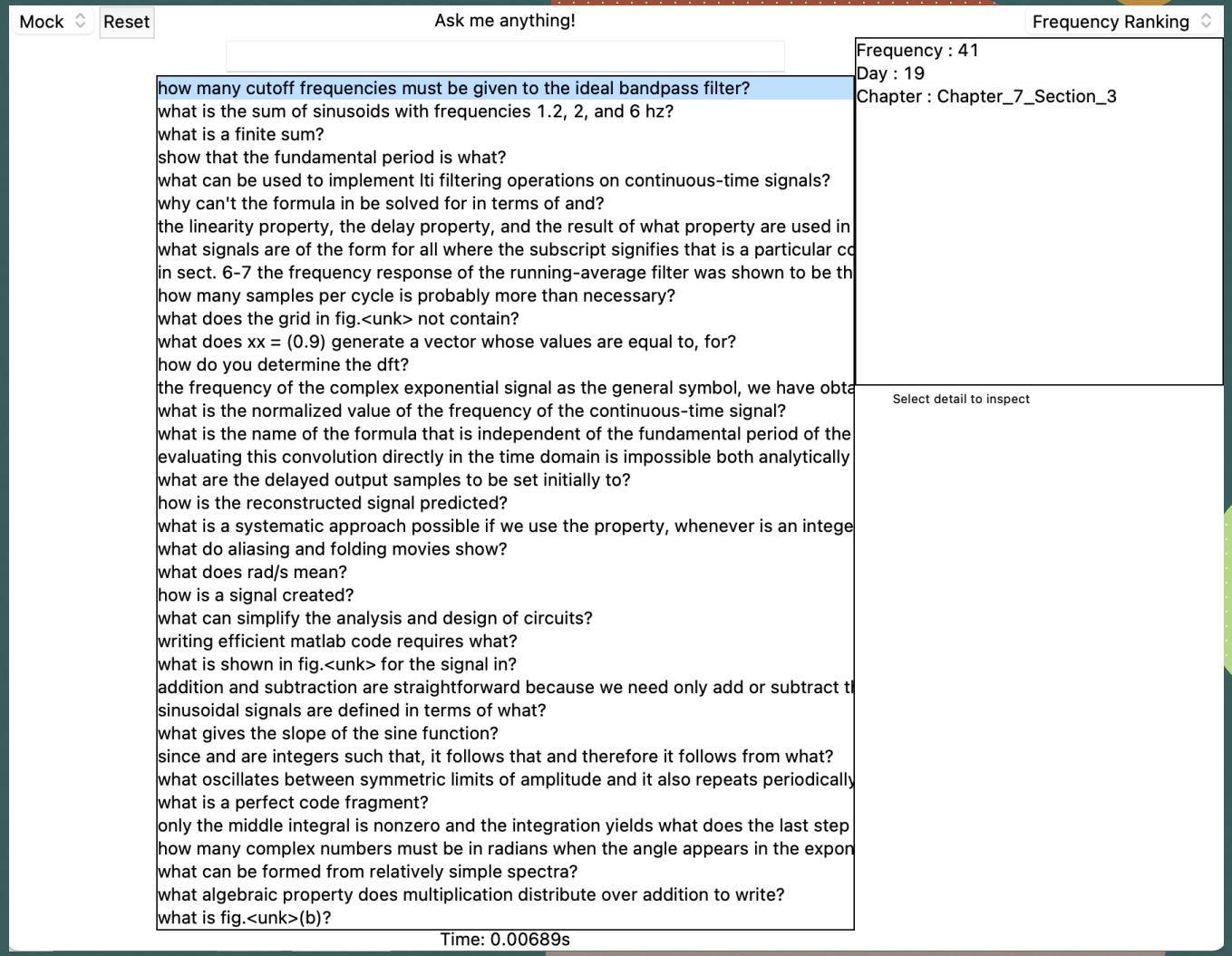
	question	chapter	frequency	time_stamp
0	what can be started with the command demo?	Chapter_-1_Section_0	1	82
1	what is the command prompt indicated by in the...	Chapter_-1_Section_0	1	8
2	what command returns text information in the c...	Chapter_-1_Section_0	3	112
3	help punct summarizes punctuation as used in w...	Chapter_-1_Section_0	1	49
4	the help system now has what?	Chapter_-1_Section_0	1	55
...	...	...	...	...
5906	what is a movie that animates the relationship...	Chapter_9_Section_9	2	94
5907	what movie animates the relationship between t...	Chapter_9_Section_9	1	47
5908	how many different filters are demonstrated?	Chapter_9_Section_9	2	20
5909	what is a three-domain movie that shows how th...	Chapter_9_Section_9	1	108
5910	the website also contains some demonstrations ...	Chapter_9_Section_9	1	99

5911 rows × 4 columns



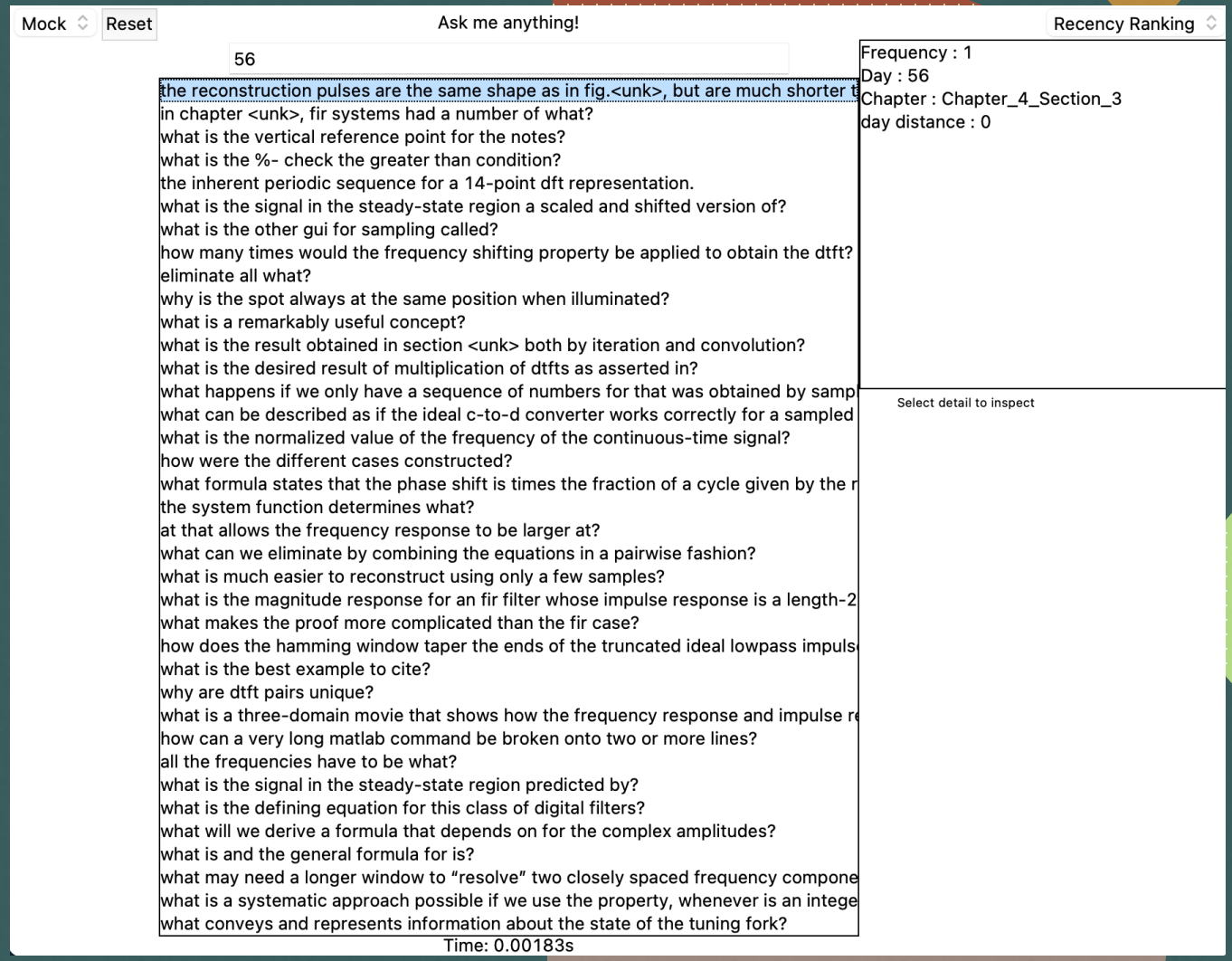
Frequency Ranking

- Rank the questions from the largest number of view to the least viewed
- Recommend most popular questions
- Can be applied for the case that no typing in the input bar



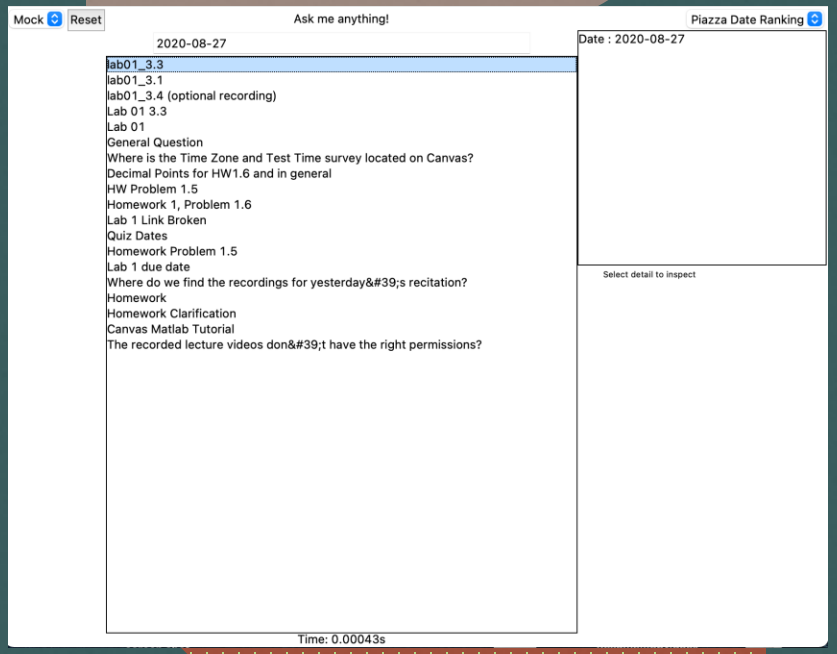
Recency Ranking

- Rank the questions according to the given day number
- Gives questions that are asked closest to the day number
- Help students check the most updated questions
- Gives an idea to rank piazza questions by dates



Piazza Date Algorithm

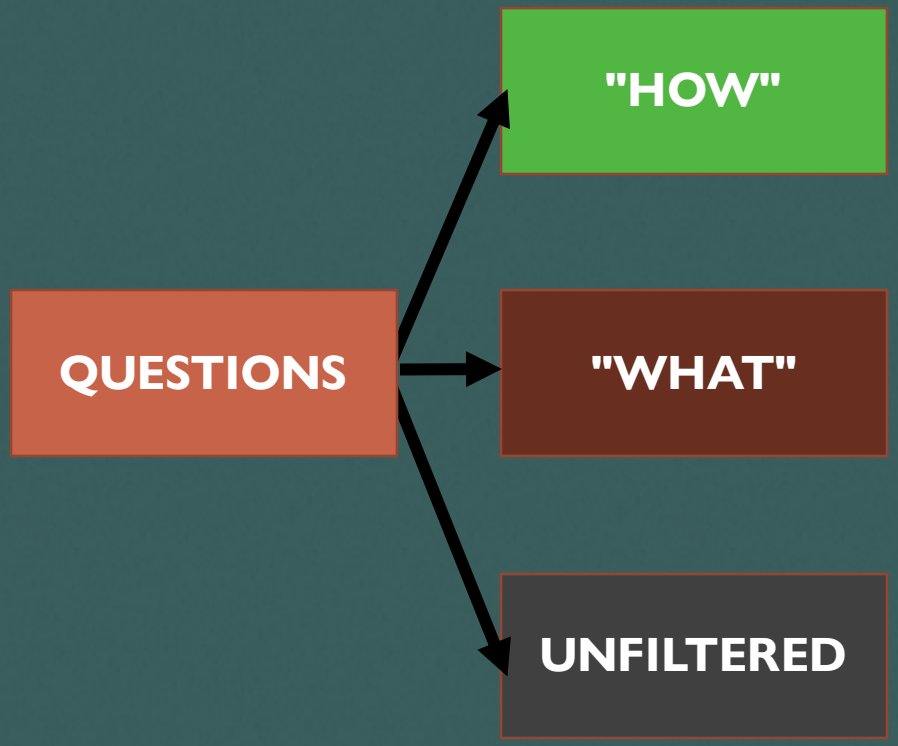
- Piazza questions from past semesters (Fall 2020) are tracked and organized into a dictionary that corresponds each question to the date they were posted
- This algorithm takes the current date and the previous Piazza data to apply a filter to the current recommended questions
  - First, the algorithm takes the current date and calculates the data that corresponds to it in each of the previous semesters
  - Then, the algorithm displays only the previous questions that are within 10 days of the new calculated date



Piazza Date Algorithm

- This algorithm is meant to prevent people from repeatedly asking the same questions in Piazza while giving the most relevant questions
- Each semester is likely to follow a similar schedule and students are likely to be learning the same material in the same part of the semester
- Advantages:
  - The algorithm would suggest questions that are likely to be like the questions that students would ask at that time
- Disadvantages:
  - Many of the questions that students ask involve previous homework and labs that are likely to change each semester

How-What Filter



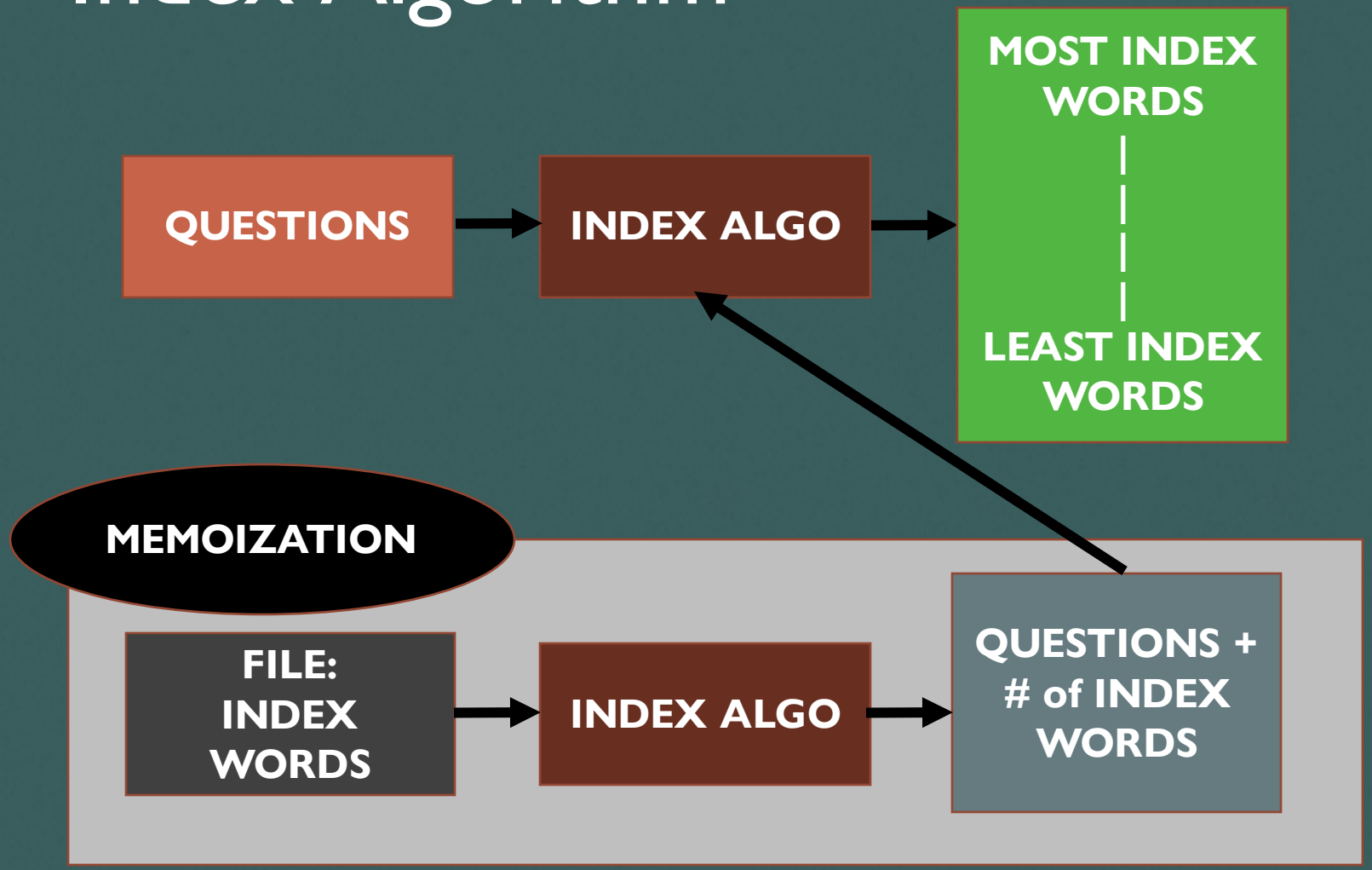
Ask me anything!

n

notice how much smaller is than?
negative frequency components should be included even though they are conjugates of v
noncausal systems cannot be implemented in a real-time application because what is not
note that the samples are connected by what?
notice that the aa vector has for its second element?
notice that the algebraic sign of the phase angles of the folded aliases must be opposite t
notice that there are only distinct roots?
notice that there are only what?
notice the alias images of the original spectrum on either side of the base band where all
notice the alias images of the original spectrum on either side of the base band?
numerical convolution of finite length signals via synthetic polynomial multiplication?

- **HOW Questions:** Procedure, critical thinking, problem-solving
- **WHAT Questions:** Objective, facts, details

Index Algorithm



MORE INDEX WORDS

RELEVANT TO STUDENT

Reset

Ask me anything!

IndexAlgo

how can we

how can we associate the property of linear phase with time delay?
how can we write a discrete fourier series for a periodic discrete-time signal?
how can we add or subtract integer multiples of from the angle of a complex number w
how can we derive a simple formula for the magnitude and phase of the averager using
how can we easily show that the overall impulse response of the other cascade system
how can we look for a connection between the poles and zeros and the shape of the fr
how can we reduce to two separate real formulas for the magnitude and phase as func
how can we represent a cosine signal as the sum of two suddenly applied complex exp
how can we show that the signal is an alias of the signal?
how can we build better a to d converters by incorporating a lowpass filter?
how can we change the discrete-time system to vary the overall response?
how can we compute the transform of an infinite-duration signal?
how can we compute the values of a discrete-time signal?
how can we control the frequency of a sinusoidal oscillator by adjusting the coefficient
how can we create a signal with quadratic angle function?
how can we find by now using the superposition property of the -transform?
how can we gain insight from the system function?
how can we multiply both sides by the complex exponential and integrate over the peri
how can we rewrite the inverse dtft integral with limits that go from 0 to?
how can we synthesize any periodic signal?
how can we analyze a signal mathematically?
how can we compare the approximations of triangle waves?
how can we compute the -transform of the convolution?
how can we compute the cartesian coordinates from the polar variables?
how can we construct a general formula for the impulse response?
how can we convert dft coefficients into a dfs representation?
how can we convert these dft coefficients into a dfs representation?
how can we define the -transform system function?

chapter : Chapter_6_Section_4
answer : in all filters
answer_start : 452
answer_end : 452
context : Delay System The delay system is a
index_word_count : 4

Delay System The delay system is a simple FIR filter given by t he difference equation It has only one nonzero filter coefficien t,, so its frequency response is For this filter, a plot of the freq uency response is easy to visualize; the magnitude response i s one for all frequencies and the phase is given by the equatio n of a straight line with a slope equal to, as in Fig.~※. As a resu lt, we can associate the property of linear phase with time del ay in all filters. Since time delay affects only the time origin of the signal in a predictable way, we often think of linear phase as an ideal phase response.

Similarity Analyzer

- Utilized 2 Data Sets to analyze Piazza question similarity with the Textbook questions
  - Piazza Data from Fall 2020
  - Stanford Textbook Questions
- Using Python's Doc2Vec library, vectorized each of the Piazza questions and the textbook questions to then compare them against one another. It also creates a model document to compare to
  - In this case, the Textbook Questions are the model document
- By vectorizing, it goes beyond just pattern matching and analyzes string semantics and more
- From there, the algorithm reports 3 most similar questions from the model document. It reports a series of indices in the model document paired with the similarity reported as a decimal
- To present it onto our GUI, the algorithm includes a "cleaning up" part that doesn't include a question if it is reported as less than 73% similar

Usage of Similarity Analyzer

- Students can use find quicker answers to their questions that they would normally post on Piazza
- Advantages:
  - For more on specific questions directly relating to course concepts, there can be very similar questions
- Disadvantages:
  - Algorithm is much slower than other algorithms
  - Some questions that have even a little bit of course content in them have trouble finding strong matches

Mock Reset

Ask me anything!

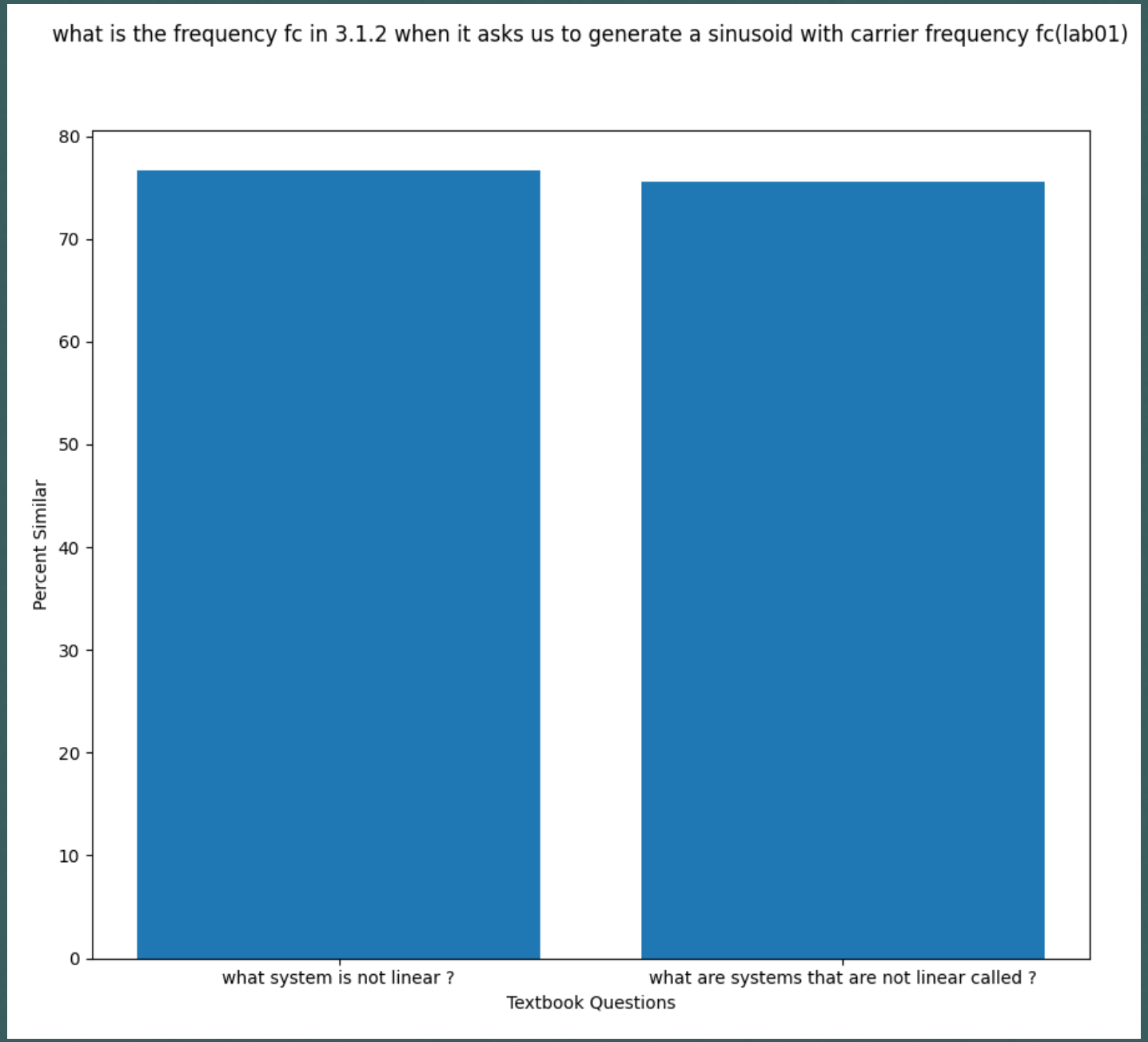
Similarity

Please follow the following guidelines when posting on Piazza:1. Piazza is a place to
Is the quiz 1 going to be during class time? Or could it be accessed through a certain
What is CanvasID on practice exam? is it same as GTID?
When I tried to watch the recorded videos M1-4, it took me to a separate site and dis
How well made does the lab report need to be? Do we need annotations/captions for
I'm confused on why we were given the phase in section 3.1 of the lab. The Mat
For lab 2, in section 3.2, there is a line that states if(nargin < 4) %-- Allow optior
The lab instructions say to attach the output wav file to the report. I was wondering h
Im having a really hard time figuring out where to start with this problem. This is the c
When is Lab 1 due? I can only see the assigned date on the file which says August 24
I understand that for the lab report, we are asked to include both plots and codes. Bu
How exactly should I study for the upcoming exam? Are there past papers or even so
For the homework's, do you want us to draw or write down the question and the
Hello, for homework 1, problem 1, it has parts a and b. I know that we have to write o
1) In question 2.3, part a, HW2, we have a negative DC value, so when representing
Quiz 1 open note, open book. Calculators are allowed. MATLAB is allowed.
what chapters will be covered this exam
Is there Fourier series on the test?
the question was write y(t), but in the provided box where to write answer,it starts wi
How do you start on these problems? I don't understand what the intuition is fo
If we are adding two polar forms on calculator and writing down the answer in the pol
I'm having trouble with plotting ece2026lab01.wav from t=0.2 to t=0.42. How c
When will the HW2 file be posted on the Canvas page?
I have a question about plotting. So I understand that the the signal should be 3 seco
For problem 2.5 in homework 2, I got the answer for A and phi, but I'm not sure
Will we lose points for not having comments in our code that describe what the code
For part b of problem 3.5, do I have to make a spectrum for when A = 2, or do I have
On the announcement it says to plot the reversed sound wave xhr from t = 0.2s to 0.
If we have an equation for a sinusoid that had to be plotted on a frequency spectrum,
Since I am outside the US, I have to print on A4 sheets (8.3 x 11.7 inches) which has
So for the end period of 2 pm , do we need to finish the test at 2pm, or we can start a
Canvas says it has 7 questions but thhere are only 6 on the template?
For Problem 1.5, do I have to account for 3 and -3 case? Or do I keep it as sqrt(9) an
When it says the real part for the expression given, do I take the real part and then ac
I am not exactly sure how to combine s(t) with the sinusoid I create. Do I just multiply
How specific do I have to be with my plots, they're just rough sketches right I don't n
Hi! In the lab 0 instruction, it mentions there are Matlab tutorials on canvas page. I tri

Time: 3e-05s

Closest Textbook Questions : ['if matrix-vect
Similarity : [82.58255124092102, 74.90602
Click Here to Visualize! :

Select detail to inspect


what is the frequency fc in 3.1.2 when it asks us to generate a sinusoid with carrier frequency fc(lab01)
80
70
60
50
40
30
20
10
0
Percent Similar
what system is not linear ?
what are systems that are not linear called ?
Textbook Questions

Demo

Further Advancements

- We have several important algorithms, but would like to have one all encompassing algorithm that combines them altogether
- We can always bring in more course resources such as syllabi to connect provide even more metrics