# 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.

### Q autoc

- Q autocracy
- Q autoclicker
- Q autocad
- Q autocratic
- autocracy definition
- Q autoclave
- Q autocrat
- autocheck
- autoclicker download
- Q autocratic government

### 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

## Al 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

Mock C Reset

#### **Autocomplete Demo**

Ask me anything!

#### Signals

many signals originate as what? how many signals can we convolve? what do most signals originate as? functions serve as what for signals? most real-world signals exhibit what? what is one application of fm signals? what are signals of this class called? lab: #4 synthesis of sinusoidal signals how can periodic signals be synthesized? how many signals are shown in fig.<unk>? what should the sum of the two signals be? how can real-world signals be represented? where are signals generated and processed? how many signals are shown in fig.<unk>(b)? what are the linear fm signals also called? how many signals does fig.<unk>(b) perform? what can we obtain discrete-time signals in? why are the following signals equal for all? what should be the phase of the two signals? what are signals that are audible to humans? what are examples of continuous-time signals? what do systems operate on signals to produce? what are the most basic signals in the universe? sinusoidal signals are defined in terms of what? what are complex exponential signals related to? what can be written for continuous time signals? what is a convolution between two signals called? how does it represent signals as well as systems? what is often not true for discrete-time signals? good examples of indefinite long signals are what? what can be used to convolve the following signals? what is an example of a nulling signals with zeros? what can remove signals that have the special form? the proof would be valid for signals of what length? what can nonperiodic signals also be represented as?

#### count\_id : 6394 question : many signals originate as what? title : Chapter\_1\_Section\_0 book\_id : 74n9jr57940eijx5oxpk1x2v5 answer : continuous-time signals answer\_start : 1710 answer\_end : 1733 context : Mathematical Representation of Si

frequency:1

SQL

Mathematical Representation of Signals Signals are patterns of variations that represent or encode information. They hav e a central role in measurement, in probing other physical sy stems, in medical technology, and in telecommunication, to n ame just a few areas. Many signals are naturally thought of a s a pattern of variations in time. A good example is a speech signal, which initially arises as a pattern of changing air pres sure in the vocal tract. This pattern, of course, evolves with t ime, creating what we often call a time waveform. Figure~\* Plot of part of a speech signal. This signal can be represente d as a function of a single (time) variable,. The shaded regio n is shown in more detail in Fig.~\*. shows a plot of a speech waveform. In this plot, the vertical axis represents air pressu re, or microphone voltage, and the horizontal axis represents time. Notice that there are four plots in the figure correspon ding to four contiguous time segments of the speech wavefo rm. The second plot is a continuation of the first, and so on, with each graph corresponding to a time interval of 50 millis econds (msec). The speech signal in Fig.~\* is an example of a one-dimensional continuous-time signal. Such signals can be represented mathematically as a function of a single inde pendent variable, which is normally denoted. Although in this

# 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?
  - <u>What d</u>ictates Newtons Second Law?
- Extremely fast, but narrow results.

#### Ask me anything!

#### How are

how are beat notes produced? how are ideal filters given? how are sample points shown? how are systems often depicted? how are filter coefficients given? how are values held in memory locations?



# SQL Pattern Matching

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

#### Ask me anything!

FIR filters

fir filters are both linear and what? what are fir filters? what can fir filters modify? what is the simplest of fir filters? why can't ideal filters be fir filters? what is a final example of fir filters? what is of interest to us for fir filters? what is a crucial building block of fir filters? what property are most fir filters designed with? in contrast to fir filters, iir filters involve what? a filter design method must produce what for fir filters? practical fir filters ideal filters are useful but not what? what can be used to construct different iir and fir filters?

# 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

|  | Ask me anything!   |  |
|--|--|--|
| FIR filters  |  |  |
| fir filters are both linear and  | id what?   |  |
| what are fir filters?<br>what can fir filters modify?<br>what is the simplest of fir fi<br>why can't ideal filters be fin<br>what is a final example of f<br>what is of interest to us for<br>what is a crucial building bi<br>what property are most fir<br>in contrast to fir filters, iir f<br>a filter design method mus<br>practical fir filters ideal filter | filters?<br>fir filters?<br>fir filters?<br>plock of fir filters?<br>filters designed with?<br>filters involve what?<br>st produce what for fir filters? |  |

what can be used to construct different iir and fir filters? what window is particularly noticeable for practical fir filters?

### 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?     | Chapter1_Section_0  | 1         | 82         |
| 1       | what is the command prompt indicated by in the | Chapter1_Section_0  | 1         | 8          |
| 2       | what command returns text information in the c | Chapter1_Section_0  | 3         | 112        |
| 3       | help punct summarizes punctuation as used in w | Chapter1_Section_0  | 1         | 49         |
| 4       | the help system now has what?                  | Chapter1_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 ro | ws × 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

| Mock 🗘 | Reset | Ask me anything!  |                        | Frequency Ranking 💲 |
|--------|-------|---|------------------------|---------------------|
|        |       |   | Frequency : 41         |                     |
|        |       | how many outoff from anoing must be given to the ideal handness filter?                 | Day : 19               |                     |
|        |       | now many cuton frequencies must be given to the ideal bandpass lifter?                  | Chapter : Chapter_7    | '_Section_3         |
|        |       | what is a finite sum?   |                        |                     |
|        |       | what is a limite sum:   |                        |                     |
|        |       | what can be used to implement it filtering operations on continuous-time signals?       | ,                      |                     |
|        |       | why can't the formula in he solved for in terms of and?                                 |                        |                     |
|        |       | the linearity property, the delay property, and the result of what property are used    | Lin                    |                     |
|        |       | what signals are of the form for all where the subscript signifies that is a particular | r cc                   |                     |
|        |       | in sect. 6-7 the frequency response of the running-average filter was shown to be       | e th                   |                     |
|        |       | how many samples per cycle is probably more than necessary?                             |                        |                     |
|        |       | what does the grid in fig. <unk> not contain?</unk>                                     |                        |                     |
|        |       | what does $xx = (0.9)$ generate a vector whose values are equal to, for?                |                        |                     |
|        |       | how do you determine the dft?   |                        |                     |
|        |       | the frequency of the complex exponential signal as the general symbol, we have c        | bta                    | - 1                 |
|        |       | what is the normalized value of the frequency of the continuous-time signal?            | Select detail to inspe | Ct                  |
|        |       | what is the name of the formula that is independent of the fundamental period of t      | the                    |                     |
|        |       | evaluating this convolution directly in the time domain is impossible both analytica    | ally                   |                     |
|        |       | what are the delayed output samples to be set initially to?                             |                        |                     |
|        |       | how is the reconstructed signal predicted?  |                        |                     |
|        |       | what is a systematic approach possible if we use the property, whenever is an inte      | ege                    |                     |
|        |       | what do aliasing and folding movies show?   |                        |                     |
|        |       | what does rad/s mean?   |                        |                     |
|        |       | how is a signal created?  |                        |                     |
|        |       | what can simplify the analysis and design of circuits?                                  |                        |                     |
|        |       | writing efficient matlab code requires what?  |                        |                     |
|        |       | what is shown in fig. <unk> for the signal in?</unk>                                    |                        |                     |
|        |       | addition and subtraction are straightforward because we need only add or subtrac        | ct ti                  |                     |
|        |       | sinusoidal signals are defined in terms of what?  |                        |                     |
|        |       | what gives the slope of the sine function?  |                        |                     |
|        |       | since and are integers such that, it follows that and therefore it follows from what    |                        |                     |
|        |       | what oscillates between symmetric limits of amplitude and it also repeats periodic      | any                    |                     |
|        |       | what is a perfect code fragment?  |                        |                     |
|        |       | how many complex numbers must be in radians when the angle appears in the even          |                        |                     |
|        |       | what can be formed from relatively simple spectra?                                      |                        |                     |
|        |       | what algebraic property does multiplication distribute over addition to write?          |                        |                     |
|        |       | what is fig <unk>(b)?</unk>   |                        |                     |
|        |       | Time: 0.00689s  | ]                      |                     |

# Recency Ranking

Mock 🗘

- 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

| Reset | Ask me anything!   | Recency Ranking 💲             |
|-------|--|-------------------------------|
|       | 56   | Frequency : 1                 |
|       |  | Day : 56                      |
|       | the reconstruction pulses are the same shape as in fig. <unk>, but are much shorter t</unk>  | Chapter : Chapter_4_Section_3 |
|       | in chapter <unk>, fir systems had a number of what?</unk>  | day distance : 0              |
|       | what is the vertical reference point for the notes?  |                               |
|       | what is the %- check the greater than condition?   |                               |
|       | the innerent periodic sequence for a 14-point diff representation.   |                               |
|       | what is the signal in the steady-state region a scaled and shifted version of?   |                               |
|       | what is the other guilor sampling called?  |                               |
|       | now many times would the frequency shifting property be applied to obtain the difference of what?  |                               |
|       | why is the snot always at the same position when illuminated?  |                               |
|       | why is the spot always at the same position when manimated:<br>what is a remarkably useful concent?  |                               |
|       | what is the result obtained in section curks both by iteration and convolution?  |                               |
|       | what is the desired result of multiplication of dtfts as asserted in?  |                               |
|       | what is the desired result of multiplication of datas as asserted in:  |                               |
|       | what happens if we only have a sequence of numbers for that was obtained by samp<br>what can be described as if the ideal c-to-d converter works correctly for a sampled | Select detail to inspect      |
|       | what is the normalized value of the frequency of the continuous-time signal?   |                               |
|       | how were the different cases constructed?  |                               |
|       | what formula states that the phase shift is times the fraction of a cycle given by the r   |                               |
|       | the system function determines what?   |                               |
|       | at that allows the frequency response to be larger at?   |                               |
|       | what can we eliminate by combining the equations in a pairwise fashion?  |                               |
|       | what is much easier to reconstruct using only a few samples?   |                               |
|       | what is the magnitude response for an fir filter whose impulse response is a length-2  |                               |
|       | what makes the proof more complicated than the fir case?   |                               |
|       | how does the hamming window taper the ends of the truncated ideal lowpass impuls   |                               |
|       | what is the best example to cite?  |                               |
|       | why are dtft pairs unique?   |                               |
|       | what is a three-domain movie that shows how the frequency response and impulse re  |                               |
|       | how can a very long matlab command be broken onto two or more lines?   |                               |
|       | all the frequencies have to be what?   |                               |
|       | what is the signal in the steady-state region predicted by?  |                               |
|       | what is the defining equation for this class of digital filters?   |                               |
|       | what will we derive a formula that depends on for the complex amplitudes?  |                               |
|       | what is and the general formula for is?  |                               |
|       | what may need a longer window to "resolve" two closely spaced frequency compone  |                               |
|       | what is a systematic approach possible if we use the property, whenever is an intege   |                               |
|       | what conveys and represents information about the state of the tuning fork?  |                               |
|       | Time: 0.00183s   |                               |

## 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



#### Ask me anything!

#### 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 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 It, 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 ᅌ                                  |
|--------|-------|--|---|
|        |       |  | Closest Textbook Questions : ['if matrix-vect |
|        |       |  | Similarity : [82.58255124092102, 74.90602     |
|        |       | Please follow the following guidelines when posting on Piazza:1. Piazza is a place to      | Click Here to Visualize! :                    |
|        |       | 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          |   |
|        |       | 'm confused on why we were given the phase in section 3.1 of the lab. The Ma               |   |
|        |       | For lab 2, in section 3.2, there is a line that states if (nargin &It 4 ) % Allow option   |   |
|        |       | The lab instructions say to attach the output way 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 d  |   |
|        |       | 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         |   |
|        |       | ⊩or the nomework'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 d     | Select detail to inspect                      |
|        |       | 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      |   |
|        |       | 18#39;m having trouble with plotting ece2026lab01.wav from t=0.2 to t=0.42. How d          |   |
|        |       | 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 1'm hot 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./ 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 / 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 ad   |   |
|        |       | 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 |   |
|        |       | Lime: 3e-05s   |   |

what is the frequency fc in 3.1.2 when it asks us to generate a sinusoid with carrier frequency fc(lab01)







### 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



