Neural Speech Decoding Competition

Statistical Learning

Statistics 640 / 444 • Fall 2016 • Competition Description, Rules & Grading

Timeline:
- Contest Opens: September 13
- Progress Report I: 5:15pm October 4
- Progress Report II: 5:15pm November 1
- Contest Closes: 5:59pm November 20
- In-Class Contest Presentations: November 29 & December 1
- Final Report Due: 5:00pm December 2

Competition Website:
(Note: You must create an account on Kaggle using your @rice.edu email to access the competition.)

Contest Description:

Neural decoding seeks to reconstruct sensory stimuli or behaviors based solely on observed neural activity from neural recordings or neuroimaging. The objective of this competition is to reconstruct audible human speech based upon intracranial electrocorticography (ECoG) data from a single subject. ECoG is an invasive technology that records electrical activity from electrodes placed directly on the cortical surface of the brain. The electrodes are placed by a neurosurgeon and are typically used to monitor epileptic activity. Of the technologies available to record human neural activity, ECoG has the highest temporal resolution, making it ideal for studying the uniquely human activity of speech. Neural speech decoding is important for understanding how the brain is organized to process speech and is critical for helping those with auditory and speech deficits.

The data we will be working with comes from an experiment where a subject listened to a sentence during each of 210 trials. While the experiment involved other stimuli, we have spliced the audio of just the sentences together for use in this competition. Your task is to build a multi-task or multi-outcome regression model using the ECoG recordings as predictors to predict the audio of the speech. Here, the original audio signal was recorded at 100 Hz (10 ms intervals) and has been processed and compressed using the short-time Fourier transform to obtain power and phase information for 32 frequencies. Your goal is to predict the power spectrum of the audio signal. Note that it is important to perform neural speech decoding in the spectral domain as regression models fit in the time domain would miss the high-frequency oscillations that characterize human speech. Hence, the outcome \( \mathbf{Y} \) is a \( T \times 32 \) matrix, with \( T = 41258 \) time points in the training set (140 sentences) and \( T = 20797 \) time points in the test set (70 sentences). ECoG recordings while the subject was listening to the sentences will be used as predictors. The original ECoG data consists of time series for 70 nodes or electrodes at various locations on the cortical surface. The time series was transformed to the spectral domain using standard multi-taper techniques to obtain 96 frequencies, ranging from approximately 0.95 Hz to 200 Hz; these 96 frequencies were then compressed into six frequency bands that are commonly studied in neural recordings. Thus, the predictors form a tensor array that is given as a matrix, \( \mathbf{X} \) that is \( T \times (70 \times 6) \) for \( T = 41258 \) training time points (140 sentences) and \( T = 20797 \) test time points (70 sentences). Additional
Numerical evaluation for the competition will be based on the root-mean-squared-error (RMSE) between the true power spectrum of the audio signal and your predictions. The test set is split into two sets consisting of 35 sentences each to form a query set that is used for the public leaderboard and a hidden set used for the private leaderboard. The public leaderboard corresponding to results on the query set will be made available in real time throughout the competition. The winner of the competition will be determined as the team with the best RMSE on the private leaderboard, which will remain hidden throughout the competition and only revealed at the competition’s conclusion. Thus, if you overfit to the query set used to determine the public leaderboard, you may perform poorly on the hidden test set.

Beyond prediction accuracy for neural speech decoding, there are a number of interesting neuroscience questions that one can answer with this data set:

- Which nodes and / or frequencies are best for speech decoding?
- Are certain nodes / frequencies better for deciphering specific words or phonemes?
- Is speech decoding at the best nodes / frequencies a contemporaneous or time-delayed operation?
- Is speech decoding at the best nodes / frequencies a linear or non-linear operation? If it is non-linear what is the decoding function?
- Can you decipher the sentences from the test set?

Grades for the competition will be based in part on prediction accuracy, statistical innovation, and addressing these scientific questions; precise grading rubrics are outlined below.

Data for this competition has been made available by the Beauchamp Lab at Baylor College of Medicine. Note that this is a new data set that had not yet been analyzed; hence all students will be required to complete the Data Use and Sharing Agreement to participate in the competition. While this is a new data set, a similar ECoG study by the same authors can be found here:


Thank you Beauchamp Lab!

**Contest Rules**:

1. To participate in the competition, students must sign and return the Data Use and Sharing Agreement form. This is due with your Progress Report I on October 4.
2. Students may work as individuals or in pairs.
3. Individuals may merge into a team up until October 4 (Progress Report I deadline).
4. If a team consists of one student taking Stat 640 and the other taking Stat 444, the team will be graded according to the rubric for Stat 640.

5. A maximum of 2 submissions is permitted per day.

6. No outside information related to the data or other data is permitted.

7. You may use methods and software published by others. If other software packages are used, they must be available to all Rice students. All outside methods and software used must be properly cited in the final report.

8. At the request of the instructor or TAs, you will be required to submit the code used to produce your final entry. Code will be checked for hard-coding and to ensure that it produces the results of the final entry.

**Grading Overview:**
- Performance in Competition: 25%
- Innovation (640) or Interesting Findings (640 & 444): 25%
- Learning Algorithms: 25%
- Presentations, Progress & Final Reports: 25%

**Competition Winners:**
Four or more teams will be declared “winners” for three aspects of the competition:

- Best decoder (prediction winner according to the private leaderboard).
- Best statistical innovation (determined by instructor and TAs).
- Best neuroscience finding (determined by instructor and neuroscientists).

These four or more teams (at least two from 444 and two from 640) will earn an automatic A+ for the competition.

**Competition Presentations:**
Up to 15 teams will have the opportunity to give a 5 minute / 5 slide presentation during the last week of class. To be considered for a class presentation, teams must submit their 5 slides to the instructor and TAs by **5pm on Wednesday, November 23rd**. The best teams will be invited to give a class presentation; these teams will be notified by 12pm on Monday, November 28th. Teams that present their findings in class will automatically receive +5 bonus points on their competition grade. In addition to an audience choice award (voted on by classmates), the presentations will be used to determine the best statistical innovation and best neuroscience finding winners.

**Requirements for Progress Reports:**
Each progress report should be a *one page* document for each team summarizing and reflecting upon your progress thus far in the competition. The report should contain one graphic or table that summarizes your results. This includes both your internal training and predicted test error as well as the error rate for each of your submissions. (Thus, you should keep track of the performance of each base learner you fit and each submission you make.) The progress report should also address the following questions: What learning algorithms have you tried? Reflect upon the performance
of the algorithms. Which ones worked well? Which performed poorly? Why? Have you innovated (640)? If so, how? Have you found anything interesting? If so what, and what methods did you use to find this? What are the future directions in which you would like to go?

**Requirements for Final Report:**

One final competition report per team (not to exceed 6 pages for 444 or 8 pages for 640) should summarize your experience throughout the competition and reflect upon what you learned. The final report should contain the following sections:

1. **Overview.** A paragraph overview of your approach to the competition and overall performance.

2. **Base Learners.** A brief summary of all the base learners tried and how they performed. Reflect upon their performance. How did you tune each method? How did you assess the error rate of each method? What can you conclude from your results?

3. **Ensembles.** A brief summary of the ensemble building methods tried and how they performed. Reflect upon their performance. What can you conclude from your results?

4. **Model Validation & Assessment.** How did you internally assess your training and testing error? Were your error rates optimistic or pessimistic for the leaderboard? Were your error rates different between the private and public test sets? Reflect upon your findings.

5. **Best Scoring Model.** Describe in detail your best performing method in terms of prediction error.

6. **Statistical Innovation (640 only).** How did you innovate? Be sure to place your innovation in the context of the statistics and machine learning literature.

7. **Interesting Neuroscience Findings.** Did you find anything interesting? What method did you use to find this? Reflect upon your findings.

8. **What I / We Learned.** What did you learn from the competition? What was the most challenging aspect? Was there anything unexpected? What would you do differently in the future?

Additionally, each report should contain the following:

- At least one graphic which summarizes your progress throughout the competition.
- Team acknowledgments. Acknowledge and clearly delineate the specific contributions of each team member.
- References or websites for publicly available software used.
- (640 only) Literature Cited. You should place your best model, innovations, and interesting findings in the context of the statistics and machine learning literature.
<table>
<thead>
<tr>
<th>Competition Grading Rubric:</th>
<th>% Grade</th>
<th>Developing (≤ B-)</th>
<th>Progress (B - B+)</th>
<th>Competent (A - A)</th>
<th>Exemplary (A+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance in Competition</td>
<td>25%</td>
<td>Bottom 25%</td>
<td>25% - 50%</td>
<td>50% - 90%</td>
<td>top 10%</td>
</tr>
<tr>
<td>Statistical Innovation (640 only)</td>
<td>10%</td>
<td>Used standard, black-box methods and techniques.</td>
<td>Trivial innovations and straightforward extensions.</td>
<td>Innovation that make an impact on overall performance.</td>
<td>Publication worthy innovation.</td>
</tr>
<tr>
<td>Interesting Findings</td>
<td>15% (640) or 25% (444)</td>
<td>Trivial inference and observations.</td>
<td>Findings that could be found by trivial analysis or provide simple insight into the data.</td>
<td>Findings that are interesting and provide new insight into the data.</td>
<td>Publication worthy findings.</td>
</tr>
<tr>
<td>Base Learners</td>
<td>10%</td>
<td>Limited set of base learners fit.</td>
<td>Fit several base learners but did not go beyond black-box methods.</td>
<td>Fit several base learners and went beyond black-box methods.</td>
<td>Fit several base learners in a way that substantially improved performance.</td>
</tr>
</tbody>
</table>

**Examples of Statistical Innovation:** Developed new statistical learning methods or algorithms; Extended existing methods or algorithms to improve performance; Built an ensemble in a novel manner; Combined methods or algorithms in a new way; Developed new computational strategies for fitting algorithms; Developed novel processing techniques that improved performance.

**Research Computing Accounts on DAVinCI & NOTS:**
To support the computing needs of this competition, each student registered for the course has been given an account on DAVinCI and NOTS, which run a Linux operating system. Information on how to log in to and work on DAVinCI and NOTS can be found here: http://www.rcsg.rice.edu/. If you are auditing the class and would like an account on DAVinCI and NOTS, please contact the TAs. Thank you Rice Center for Research Computing!

Thank you Kaggle In Class!