NEU502B Homework 4: Multivariate pattern analysis

Due March 27, 2024

Submission instructions: First, rename your homework notebook to include your name (e.g. `homework-4-nastase.ipynb`); keep your homework notebook in the `homework` directory of your clone of the class repository. Prior to submitting, restart the kernel and run all cells (see *Kernel > Restart Kernel and Run All Cells...*) to make sure your code runs and the figures render properly. Only include cells with necessary code or answers; don't include extra cells used for troubleshooting. To submit, `git add`, `git commit`, and `git push` your homework to your fork of the class repository, then make a pull request on GitHub to sync your homework into the class repository.

In this homework assignment, you will work through three commonly used methods in cognitive computational neuroscience: (1) neural decoding via multivariate pattern analysis (MVPA); (2) representational similarity analysis (RSA); and (3) voxelwise encoding analysis using regularized regression. Each of these problems builds on tools and ideas we've introduced in the in-class lab notebooks.

In [1]:
```
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import pickle, os
%autosave 60
```

```
Autosaving every 60 seconds
```

Problem 1: Multivariate pattern classification

First, we'll start with a simple example of classifying distributed response patterns for different object categories from Haxby et al., 2001 (https://doi.org/10.1126/science.1063736). We'll begin by loading in the data, as well as labels for the stimuli and runs. You'll need to change `data_dir` to a directory on your computer (or the server); if you've already downloaded this dataset in lab, you can set `data_dir` to the existing directory to save time.

In [2]:
```
from nilearn import datasets
from nilearn.image import index_img
import pandas as pd

# Change this path to a directory on your computer!
data_dir = '/scratch/gpfs/al8996/nilearn_data'

# Load the Haxby et al., 2001 data via Nilearn
haxby_dataset = datasets.fetch_haxby(data_dir=data_dir)

# Load in session metadata as pandas DataFrame
session = pd.read_csv(haxby_dataset.session_target[0], sep=" ")

# Extract stimuli and run labels for this subject
stimuli, runs = session['labels'].values, session['chunks'].values

# Create a boolean array indexing TRs containing a stimulus (non-rest)
task_trs = stimuli != 'rest'

# Get list of unique stimulus categories (excluding rest)
categories = [c for c in np.unique(stimuli) if c != 'rest']

# Extract task TRs for fMRI data and stimulus/run labels
func_task = index_img(haxby_dataset.func[0], task_trs)
stimuli_task = stimuli[task_trs]
runs_task = runs[task_trs]
```

Use `NiftiMasker` (with `standardize=True`) to create a masker for ventral temporal (VT) cortex. Use the masker to extract the the NumPy array containing the functional data. (We'll analyze the data using scikit-learn rather than nilearn.)

In [3]:
```
# Get the VT mask file and creater masker:
from nilearn.maskers import NiftiMasker
masker_vt = NiftiMasker(mask_img=haxby_dataset.mask_vt[0], standardize=T

# Uses masker to extract numpy array for VT:
func_data = masker_vt.fit_transform(func_task)
print(f"Masked data shape: {func_data.shape}")
```

```
/home/al8996/.conda/envs/502b/lib/python3.12/site-packages/nilearn/imag
e/resampling.py:492: UserWarning: The provided image has no sform in it
s header. Please check the provided file. Results may not be as expecte
d.
  warnings.warn(

Masked data shape: (864, 464)
```

Now, we'll set up a full SVM classification analysis using leave-one-run-out outer cross-validation with a nested leave-one-run-out inner cross-validation loop for grid search across the values of the SVM regularization parameter C. Sounds like a lot! But scikit-learn makes it pretty straightforward. First, initialize the `LinearSVC` estimator. Since this well-behaved dataset has the same number of samples for each stimulus category in each run, we can perform leave-one-run-out cross-validation using just `KFold` rather than having to specify the runs directly. Initalize an outer `KFold` cross-validator with 12 splits and an inner `KFold` cross-validator

with 11 splits. We'll search over a handful of C parameters: `param_grid = {'C': [1e-2, 1e-1, 1]}` . Initialize the `GridSearchCV` estimator with the SVM estimator, the parameter grid, and the inner cross-validator; then, submit this estimator to `cross_val_predict` with the outer cross-validator to run the full analysis. (This may take a few minutes to run!)

In [4]:
```
# Suppress some warnings (e.g. SVM convergence) just to clean up output
import warnings
from sklearn.model_selection import (cross_val_predict,
                                     GridSearchCV,
                                     KFold)
from sklearn.svm import LinearSVC
warnings.filterwarnings("ignore")

# Save results
outfn = 'y_pred.pkl'

# Initialize SVM and outer/inner CVs:
svm = LinearSVC()
outer_cv = KFold(n_splits=12)
inner_cv = KFold(n_splits=11)

# Set up parameter grid:
param_grid = {'C': [1e-2, 1e-1, 1]}

# Initialize GridSearchCV estimator:
grid_search = GridSearchCV(svm, param_grid, cv=inner_cv)

# Generate predictions using cross_val_predict:
if not os.path.exists(outfn):
    y_pred = cross_val_predict(grid_search, func_data, stimuli_task,
                               cv=outer_cv)
    with open(outfn, 'wb') as f:
        pickle.dump(y_pred, f)
else:
    with open(outfn, 'rb') as f:
        y_pred = pickle.load(f)
```

Inspect the resulting predictions. We'll evaluate our classifier's predictions in two ways. First, use `accuracy_score` from `sklearn.metrics` to evaluate the predictions (across all test sets) against the actual labels in terms of a single classification accuracy. Procedurally, this is slightly different from computing accuracies on the test for each fold and averaging them—but the resulting value should be the same.

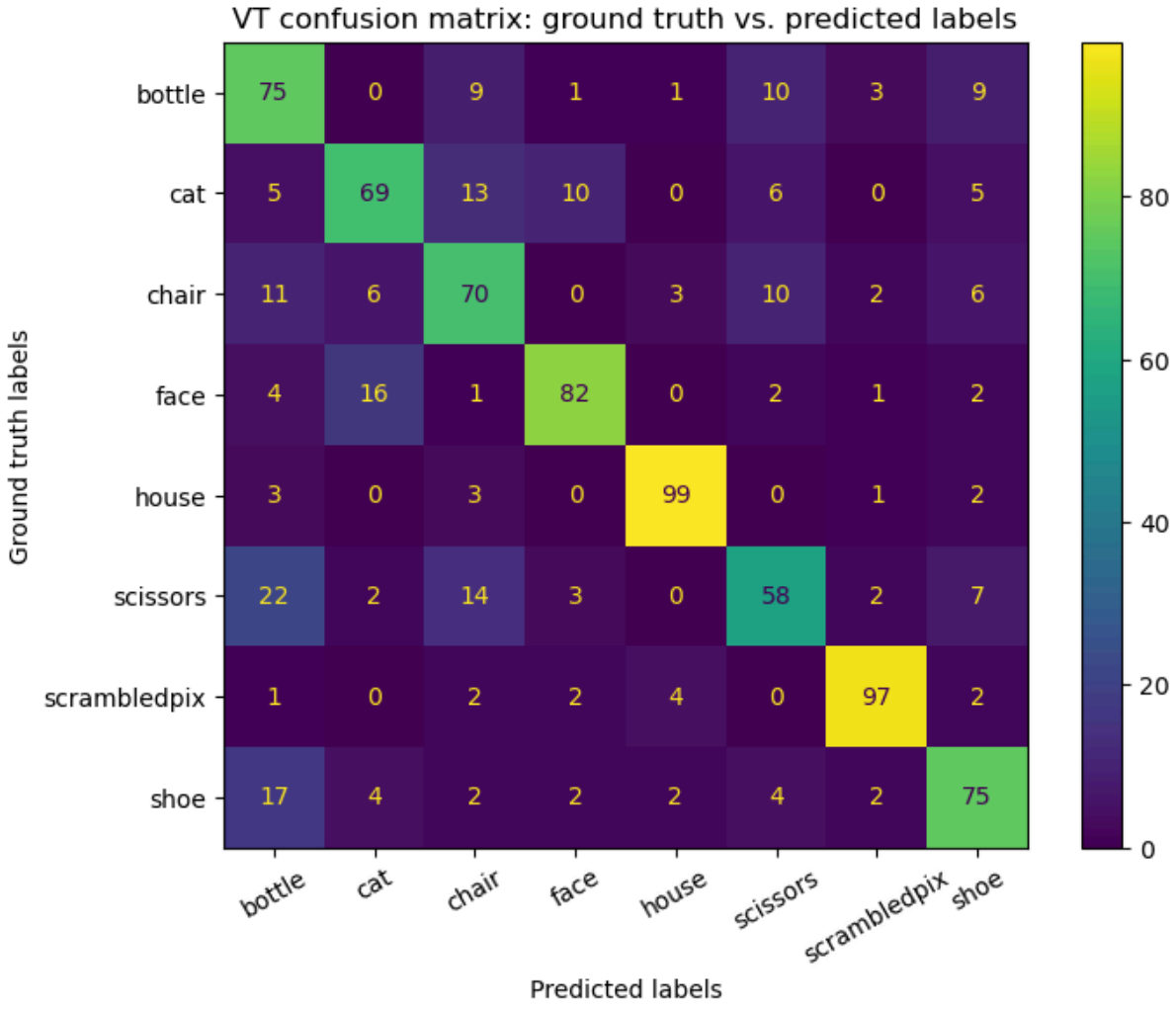
In [5]:
```
# Print accuracy score:
from sklearn.metrics import accuracy_score
accuracy = accuracy_score(stimuli_task, y_pred)
print(f"VT decoding accuracy: {np.round(accuracy, 3)}")
```

```
VT decoding accuracy: 0.723
```

To better understand what our classifer is doing (i.e. what it's getting right and what it's getting wrong), we'll construct a confusion matrix. Construct the confusion matrix from the actual stimulus labels and the classifer's predicted labels and plot it below. What categories does the

classifier tend to misclassify?

In [6]:
```
# Create confusion matrix from true and predicted labels:
from sklearn.model_selection import cross_val_predict
from sklearn.metrics import confusion_matrix, ConfusionMatrixDisplay

# Plot confusion matrix:
cm = confusion_matrix(stimuli_task, y_pred)
f, ax = plt.subplots(figsize=(8, 6))
display = ConfusionMatrixDisplay(cm, display_labels=categories)
display.plot(ax=ax)
ax.set_xticklabels(labels=ax.get_xticklabels(), rotation=30)
ax.set_ylabel("Ground truth labels")
ax.set_xlabel("Predicted labels")
ax.set_title("VT confusion matrix: ground truth vs. predicted labels")
f.show()
```



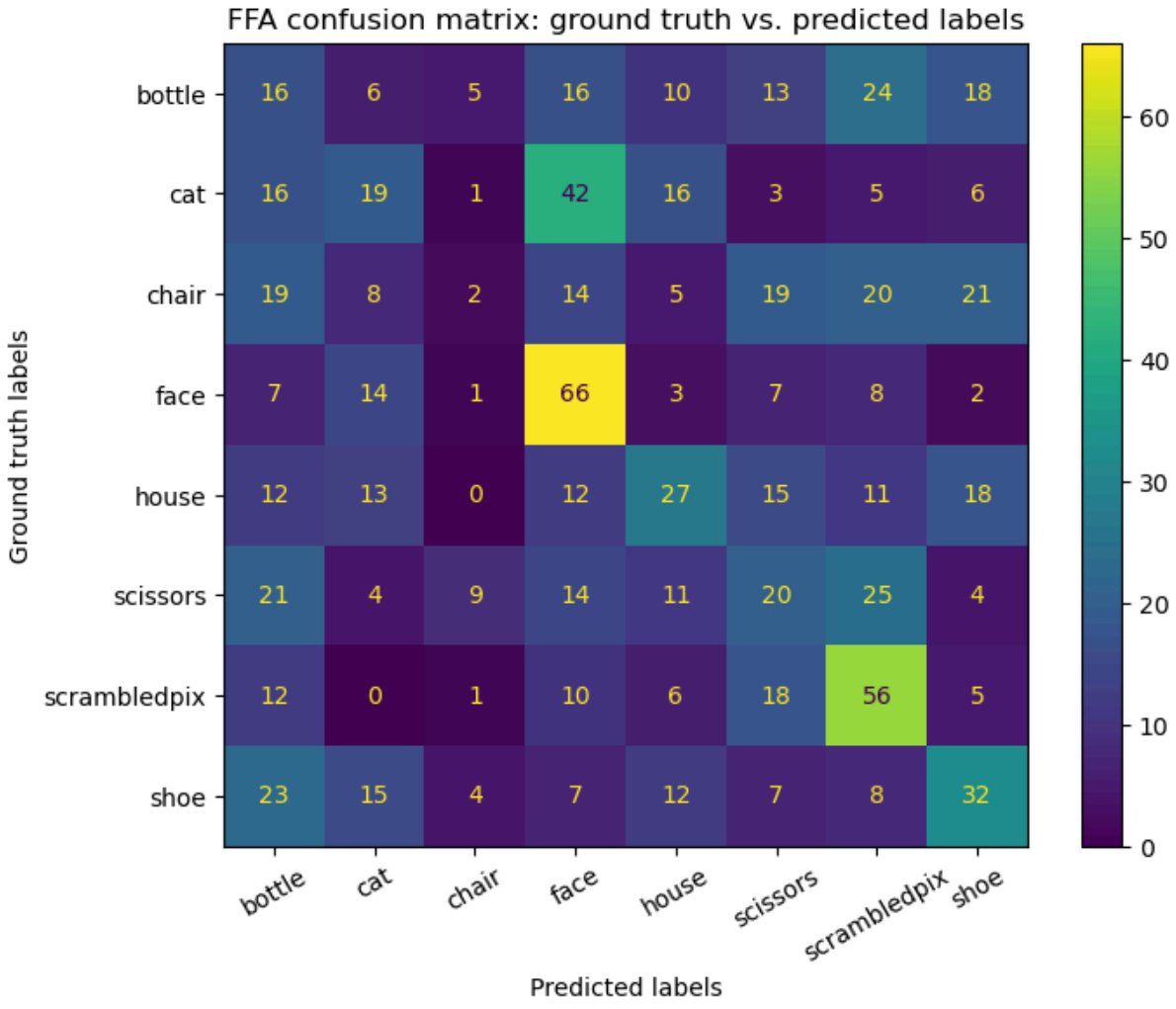
Lastly, we'll repeat the same analysis for functional regions of interest (ROIs) maximally responsive to faces (roughly FFA) and houses (roughly PPA). Use the `mask_face` and `mask_house` files from the dataset to create an FFA masker and a PPA masker; extract the functional data for both. Submit these datasets to the same analysis as above, and visualize the results in terms of an overall accuracy score and confusion matrix. Interpret the accuracies and confusion matrices in light of the expected chance accuracy, given what you know about these ROIs.

In [7]:
```
# Create masker for FFA:
masker_ffa = NiftiMasker(mask_img=haxby_dataset.mask_face[0], standardiz

# Create masker for PPA:
masker_ppa = NiftiMasker(mask_img=haxby_dataset.mask_house[0], standardi

# Uses masker to extract numpy array for ppa & ffa:
func_data_ffa = masker_ffa.fit_transform(func_task)
func_data_ppa = masker_ppa.fit_transform(func_task)
```

In [8]:
```
# Save results
outfn = 'y_pred_ffa.pkl'

# Initialize SVM and outer/inner CVs:
svm = LinearSVC()
outer_cv = KFold(n_splits=12)
inner_cv = KFold(n_splits=11)

# Set up parameter grid:
param_grid = {'C': [1e-2, 1e-1, 1]}

# Initialize GridSearchCV estimator:
grid_search = GridSearchCV(svm, param_grid, cv=inner_cv)

# Generate predictions using cross_val_predict:
if not os.path.exists(outfn):
    y_pred = cross_val_predict(grid_search, func_data_ffa, stimuli_task,
                               cv=outer_cv)
    with open(outfn, 'wb') as f:
        pickle.dump(y_pred, f)
else:
    with open(outfn, 'rb') as f:
        y_pred = pickle.load(f)
```

In [9]:
```
# Print accuracy score and plot confusion matrix:
accuracy = accuracy_score(stimuli_task, y_pred)
print(f"FFA decoding accuracy: {np.round(accuracy, 3)}")

# Plot confusion matrix:
cm = confusion_matrix(stimuli_task, y_pred)
f, ax = plt.subplots(figsize=(8, 6))
display = ConfusionMatrixDisplay(cm, display_labels=categories)
display.plot(ax=ax)
ax.set_xticklabels(labels=ax.get_xticklabels(), rotation=30)
ax.set_ylabel("Ground truth labels")
ax.set_xlabel("Predicted labels")
ax.set_title("FFA confusion matrix: ground truth vs. predicted labels")
f.show()
```

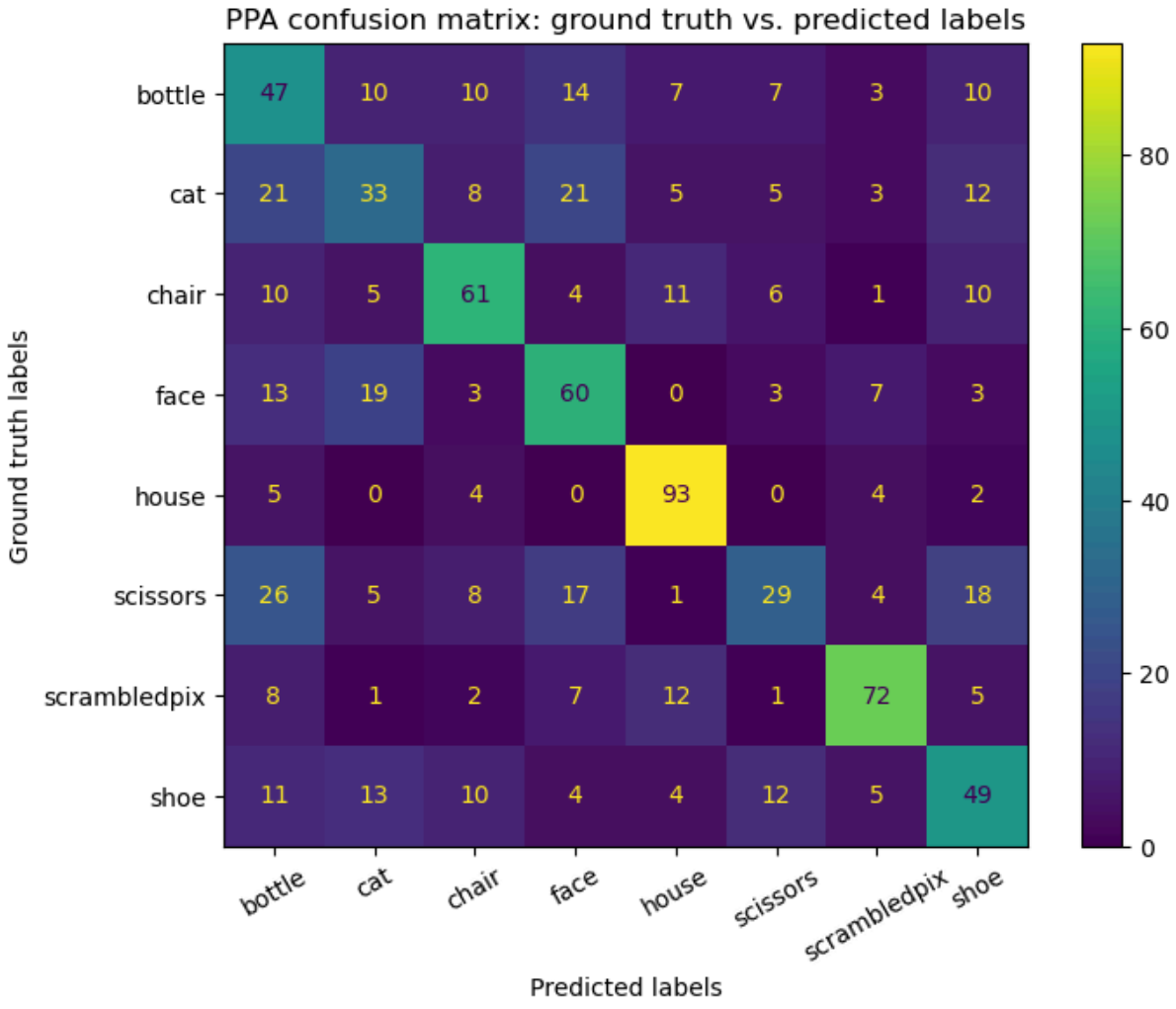
```
FFA decoding accuracy: 0.275
```



In [10]:
```
# Save results
outfn = 'y_pred_ppa.pkl'

# Initialize SVM and outer/inner CVs:
svm = LinearSVC()
outer_cv = KFold(n_splits=12)
inner_cv = KFold(n_splits=11)

# Set up parameter grid:
param_grid = {'C': [1e-2, 1e-1, 1]}

# Initialize GridSearchCV estimator:
grid_search = GridSearchCV(svm, param_grid, cv=inner_cv)

# Generate predictions using cross_val_predict:
if not os.path.exists(outfn):
    y_pred = cross_val_predict(grid_search, func_data_ppa, stimuli_task,
                               cv=outer_cv)
    with open(outfn, 'wb') as f:
        pickle.dump(y_pred, f)
else:
    with open(outfn, 'rb') as f:
        y_pred = pickle.load(f)
```

In [11]:
```
# Print accuracy score and plot confusion matrix:
accuracy = accuracy_score(stimuli_task, y_pred)
print(f"PPA decoding accuracy: {np.round(accuracy, 3)}")

# Plot confusion matrix:
cm = confusion_matrix(stimuli_task, y_pred)
f, ax = plt.subplots(figsize=(8, 6))
display = ConfusionMatrixDisplay(cm, display_labels=categories)
display.plot(ax=ax)
ax.set_xticklabels(labels=ax.get_xticklabels(), rotation=30)
ax.set_ylabel("Ground truth labels")
ax.set_xlabel("Predicted labels")
ax.set_title("PPA confusion matrix: ground truth vs. predicted labels")
f.show()
```

```
PPA decoding accuracy: 0.514
```



Problem 2: Representational similarity analysis

In this problem, we'll apply representational similarity analysis (RSA) to the human fMRI dataset from Kriegeskorte et al., 2008 (https://doi.org/10.1016/j.neuron.2008.10.043). We'll begin by loading in the ROI data and labels.

In [12]:
```
# Load in Kriegekorte dataset and labels
kriegeskorte_dataset = dict(np.load('kriegeskorte_dataset.npz',
                                    allow_pickle=True))

roi_data = kriegeskorte_dataset['roi_data'].item()
category_names = kriegeskorte_dataset['category_names']
category_labels = kriegeskorte_dataset['category_labels']
images = kriegeskorte_dataset['images']
subject_labels = ['KO', 'SN', 'TI']
roi_labels = ['lFFA', 'rFFA', 'lPPA', 'rPPA']
```

In [13]:
```
print(category_labels.shape)
print(category_names)
```

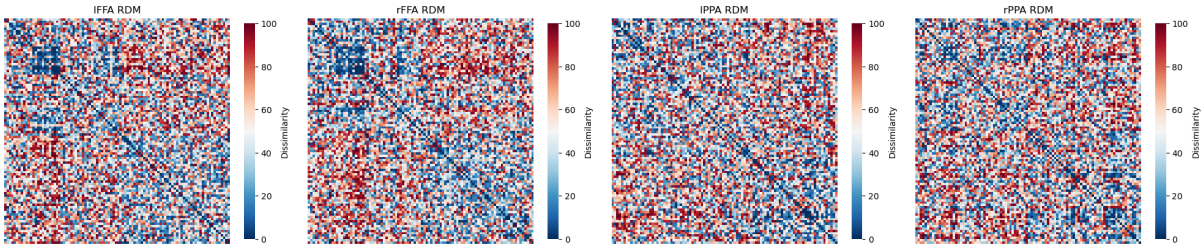
```
(96,)
['artificial inanimate' 'human bodypart' 'human face' 'natural inanimat
e'
 'nonhuman bodypart' 'nonhuman face']
```

We provide a `rank_percentile` function for visualizing RDMs in a way that more closely matches the paper.

In [14]:
```
from scipy.stats import rankdata

def rank_percentile(a):
    return rankdata(a) / len(a) * 100
```

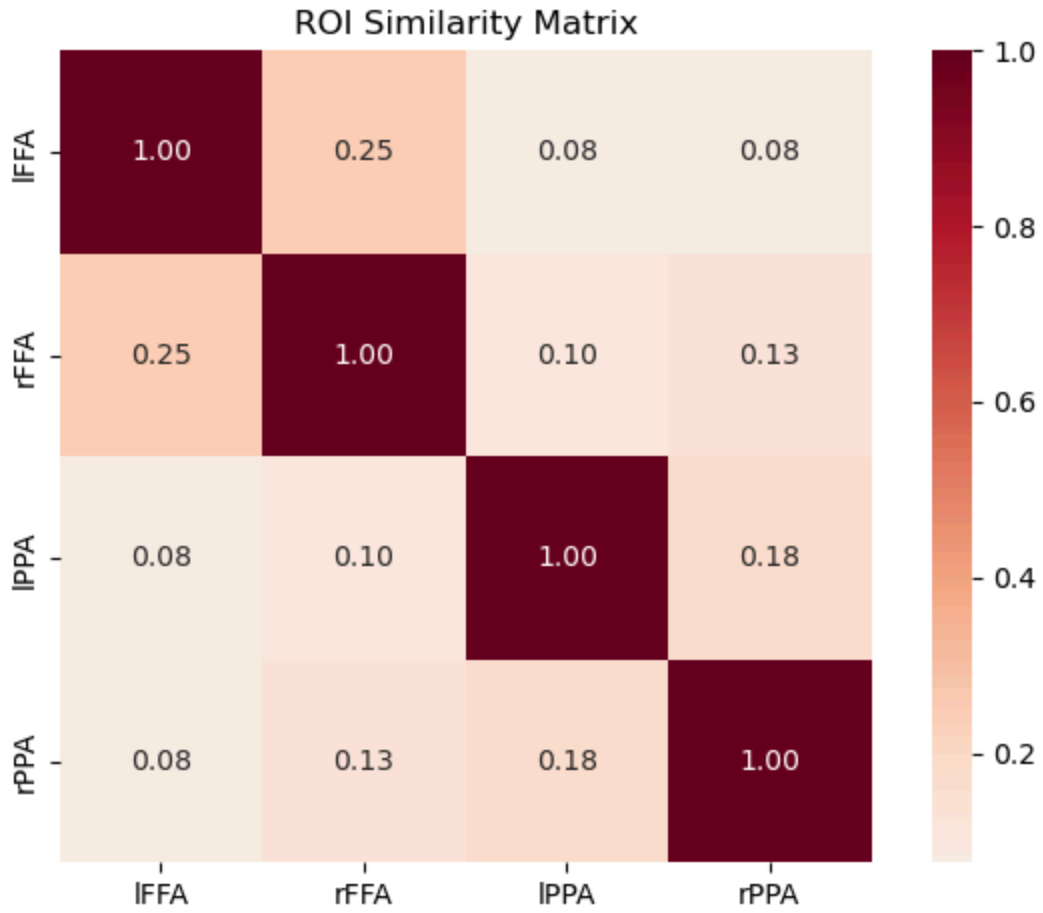
First, compute RDMs for the `'lFFA'`, `'rFFA'`, `'lPPA'`, and `'rPPA'` ROIs for subject `'TI'` using correlation distance. Here, we recommend z-soring each voxel across samples prior to computing the pairwise dissimilarities. Plot the RDMs for each ROI using the `rank_percentile` function provided above.

In [15]:
```
from scipy.stats import zscore
from scipy.spatial.distance import pdist, squareform
# Get subject data & zscore
ti_data = roi_data['TI']
ti_data_z = {roi: zscore(data, axis=0) for roi, data in ti_data.items()}

# Plot RDMs
rdms = {roi: rank_percentile(pdist(data, metric='correlation')) for roi,

f, axs = plt.subplots(1, 4, figsize=(20, 5))
for ax, roi in zip(axs, roi_labels):
    sns.heatmap(squareform(rdms[roi]),
                square=True,
                cbar_kws={'label': 'Dissimilarity', 'shrink': 0.75},
                cbar=True, cmap='RdBu_r',
                ax=ax,
                vmin=0, vmax=100,
                xticklabels=False, yticklabels=False)
    ax.set_title(f"{roi} RDM")
f.tight_layout()
f.show()
```



RSA allows us to compare the representational geometries of different ROIs. Compute the correlation between each pair of the four ROIs. Plot this similarity matrix. Which ROIs have the most similar representational geometries?

In [16]:
```
rdms_stacked = np.stack([rdms[roi] for roi in rdms.keys()])

sns.heatmap(np.corrcoef(rdms_stacked),
            annot=True,
            fmt=".2f",
            xticklabels=roi_labels,
            yticklabels=roi_labels,
            cmap='RdBu_r', center=0,
            square=True)
plt.title("ROI Similarity Matrix")
plt.tight_layout()
plt.show()
```



Stack all four ROIs to create a single combined ROI for each subject `'SN'` and `'TI'`. What is the Spearman correlation between `'SN'`'s and `'TI'`'s representational geometries?

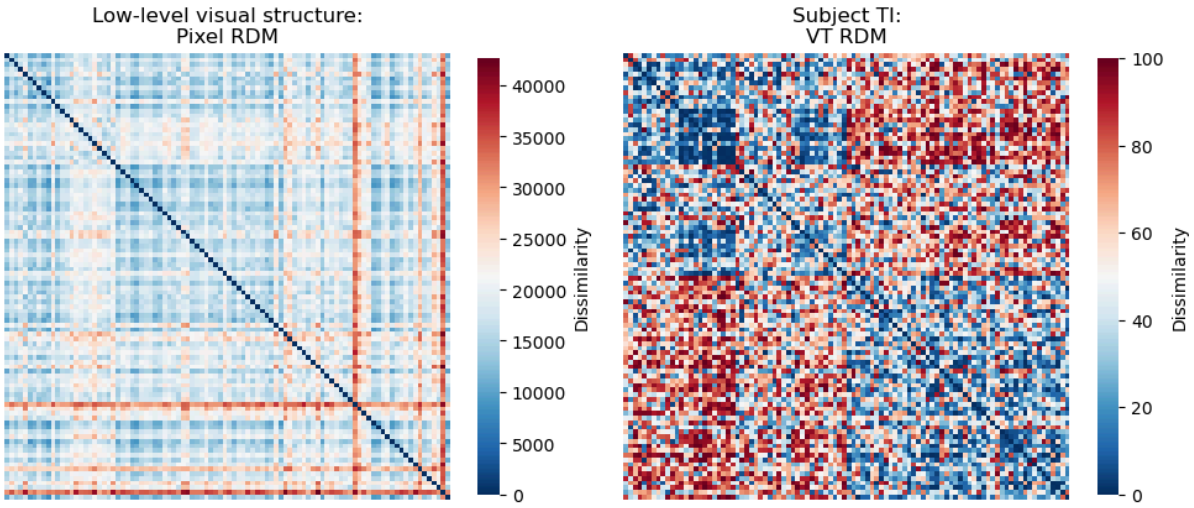
In [17]:
```
# Combine SN and TI ROIs into single VT ROI and compute RDMs:
from scipy.stats import spearmanr

# Combine SN & TI rois into a single VT ROI
vt_ti = np.column_stack([ti_data[roi] for roi in ti_data])
vt_ti = zscore(vt_ti, axis=0)

sn_data = roi_data['SN']
vt_sn = np.column_stack([sn_data[roi] for roi in sn_data])
vt_sn = zscore(vt_sn, axis=0)


# Compute correlations between SN and TI's VT RDMs:
rdm_vt_ti = rank_percentile(pdist(vt_ti, metric='correlation'))
rdm_vt_sn = rank_percentile(pdist(vt_sn, metric='correlation'))
print(f"Spearman correlation between SN & TI: {spearmanr(rdm_vt_ti, rdm_
```

```
Spearman correlation between SN & TI: 0.395
```

We can test different "model" RDMs according to how well they approximate a given neural. Here, for the sake of brevity, we'll construct an extremely simple RDM capturing low-level visual structure. Flatten each image file into a one-dimensional array of pixel values (across three color channels). Next, compute the pairwise Euclidean distances between these image vectors to construct an RDM capture low-level visual similarities. Plot this pixel RDM and compute it's Spearman correlation with `'TI'` s VT RDM?

In [18]:
```
# Create a pixel-based RDM:
images_flat = np.array([img.flatten() for img in images])
rdm_pixels = pdist(images_flat, metric='euclidean')

# Plot RDMs
rdm_list = [rdm_pixels, rdm_vt_ti]
labels = ['Low-level visual structure:\nPixel RDM', 'Subject TI:\nVT RDM
f, axs = plt.subplots(1, 2, figsize=(10, 5))
for ax, rdm, label in zip(axs, rdm_list, labels):
    sns.heatmap(squareform(rdm),
                square=True,
                cbar_kws={'label': 'Dissimilarity', 'shrink': 0.75},
                cbar=True, cmap='RdBu_r',
                ax=ax,
                xticklabels=False, yticklabels=False)
    ax.set_title(label)
f.tight_layout()
f.show()

# Compute correlations with VT RDM:
print(f"Spearman correlation between low-level visual structure RDM & sul
```

```
Spearman correlation between low-level visual structure RDM & subject T
I VT RDM: 0.022
```

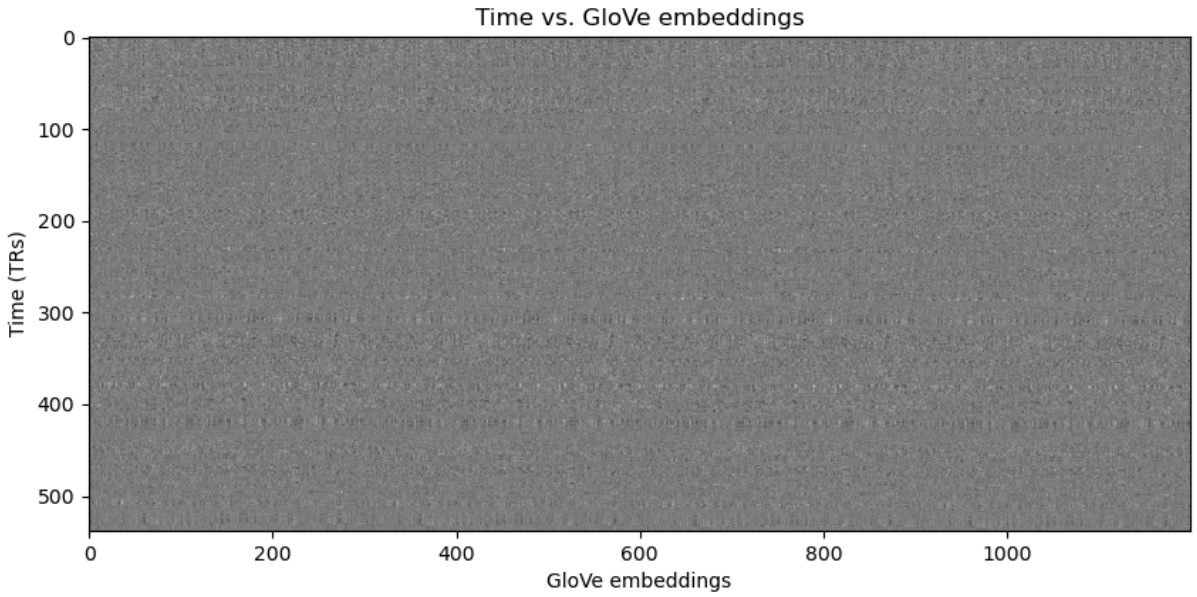


Problem 3: Voxelwise encoding analysis

In this problem, we'll return to *encoding analysis*, using regularized regression and out-of-sample prediction in individual voxels. We will use word embeddings derived from the natural language processing (NLP) model GloVe to map semantic encoding onto the brain. You can simply load the `story_transcript.txt` file in a text editor to visualize the transcript for the spoken story by Carol Daniel (https://themoth.org/stories/i-knew-you-were-black). Each line of this file corresponds to a TR in the fMRI data. Next, we extracted word embeddings from GloVe for each word in each TR. For TRs containing multiple words, we averaged the embeddings. Finally, we horizontally stacked the embeddings at lags of 2, 3, 4, and 5 TRs (3, 4.5, 6, and 7.5

In [24]:
```
# Load and visualize word embeddings:
story_embeddings = np.load("story_embeddings.npy")
print(f"Story embeddings shape: {story_embeddings.shape}")
print("538 (TRs) x 1200 (embedding dimension x number of lags)")

f, ax = plt.subplots(1, 1, figsize=(10, 10))
ax.imshow(story_embeddings, cmap='binary_r')
ax.set_xlabel("GloVe embeddings")
ax.set_ylabel("Time (TRs)")
ax.set_title("Time vs. GloVe embeddings")
f.show()
```

```
Story embeddings shape: (538, 1200)
538 (TRs) x 1200 (embedding dimension x number of lags)
```



We used fMRI to measure a subject's brain activity while they listened to the spoken story. Here, to reduce computational demands, we have spatially downsampled the fMRI data using an atlas containing 400 parcels. That is, for each parcel, we averaged the voxel time series within that parcel. Rather than fitting encoding models to tens of thousands of voxels, we'll fit our encoding model to each of the 400 parcels. Load in the `story_parcels.npy` dataset as well as the `story_atlas.nii.gz` NIfTI image from which the parcels were derived (for later visualization).

In [25]:
```
# Load in parcel time series:
story_parcels = np.load("story_parcels.npy")

# Load in the Schaefer 400-parcel atlas:
import nibabel as nib
story_atlas = nib.load('story_atlas.nii.gz')
```

Our word embedding "model" is much wider than the number of samples, so we'll need to use regularization and out-of-sample prediction to mitigate overfitting. We'll use ridge regression to fit encoding models to predict the parcel time series from the word embeddings. First, set up an split-half outer cross-validator using `KFold` with `n_splits=2` ; next, set up an inner cross-

validator using `KFold` with `n_splits=5` to perform grid search for the `alpha` hyperparameter using 5-fold cross-validation within each training set of the otuer loop. Initialize your `RidgeCV` estimator with the inner cross-validator and the following grid of alphas: `alphas = [0.1, 1.0, 10.0, 100.0, 1000.0, 10000.0]` . For each training and testing split of the other cross-validation loop, fit the ridge model on the training set of embeddings and parcel time series, and generate predicted parcel time series from the test embeddings. Compile these predicted parcel time series for model evaluation in the next step:

In [26]:
```
# Set up outer/inner cross-validators:
from sklearn.model_selection import KFold

X = story_embeddings
y = story_parcels

outer_cv = KFold(n_splits=2)
inner_cv = KFold(n_splits=5)

# Initialize RidgeCV with alpha grid and inner CV:
from sklearn.linear_model import RidgeCV
alphas = [0.1, 1.0, 10.0, 100.0, 1000.0, 10000.0]
ridge_cv = RidgeCV(alphas=alphas, cv=inner_cv, scoring='r2')
predicted_vals, gt_vals = [], []

# Loop through outer CV splits
for train_idx, test_idx in outer_cv.split(X):
    # Split data into training and testing sets
    X_train, X_test = X[train_idx], X[test_idx]
    y_train, y_test = y[train_idx], y[test_idx]

    # Fit the model on the training set
    ridge_cv.fit(X_train, y_train)

    # Generate predictions on the test set
    y_pred = ridge_cv.predict(X_test)

    # Store predictions and true values
    predicted_vals.append(y_pred)
    gt_vals.append(y_test)

predicted_vals = np.vstack(predicted_vals)
gt_vals = np.vstack(gt_vals)
```

To evaluate our encoding model's predictions, correlate the predicted parcel time series with the actual parcel time series for each parcel.

In [27]:
```
# Compute correlation between predicted and actual responses:
from scipy.stats import pearsonr
parcel_corrs = []
for i in range(story_parcels.shape[1]):
    parcel_corrs.append(spearmanr(predicted_vals[:, i], gt_vals[:, i]).c
parcel_corrs = np.array(parcel_corrs)
```

Finally, to visualize the performance of our semantic encoding model on the brain, we need to use the atlast NIfTI image to convert from parcels back to the original brain image. You can start by creating an empty brain image (i.e. zeros) the size of the atlas image. Next, loop through each parcel and insert the prediction scores (i.e. correlations between actual and predicted parcel time series) into all voxels where the atlas correponds to that parcel label. Convert this image to a NIfTI image and visualize with `plot_stat_map`; you may want to set a particular `vmax` and use a `threshold` to exclude voxels with poor prediction performance for the sake of visualization.

In [28]:
```
# Create an empty brain image and populate with parcelwise performance va
from nilearn.plotting import plot_stat_map

# Convert to NIfTI image for visualization with Nilearn:
atlas_data = story_atlas.get_fdata()
brain_img = np.zeros(story_atlas.shape)
for i, corr in enumerate(parcel_corrs):
    brain_img[atlas_data == i + 1] = corr
brain_nifti = nib.Nifti1Image(brain_img, story_atlas.affine)

# Plot correlations to visualize superior temporal cortex:
plot_stat_map(brain_nifti, cmap='RdYlBu_r', display_mode='ortho', cut_co
              threshold=0.1, vmax=0.4, title='Encoding model performance
# Plot correlations to visualize posterior medial cortex:
plot_stat_map(brain_nifti, cmap='RdYlBu_r', display_mode='ortho', cut_co
              threshold=0.1, vmax=0.4, title='Encoding model performance
```

Out[28]:
```
<nilearn.plotting.displays._slicers.OrthoSlicer at 0x14dab698d760>
```