

SAS[®] Viya[®] Natural Language Processing and Computer Vision Exam

Loading and Exploring Data (18 – 22%)

Import documents for analysis

- Convert documents for analysis.
- Explore and prepare a document.
- Troubleshoot Language encoding issues (ASCII, UTF-8, etc.).
- Given a scenario, ensure minimal loss of information when converting documents from proprietary formats to SAS supported formats.

Create and explore a project in SAS Visual Text Analytics

- Identify the SAS Visual Text Analytics default pipeline.
- Explore the Documents Table.
- Identify and define key features of the term table.
- Given a scenario, appropriately assign text and category roles.
- Export score code to score new data sets.

Load and prepare image data

- Load labeled image data (labelLevels, loadImages).
- Augment image data.
- Prepare data for modeling.

Identifying Text Patterns Using Natural Language Processing Techniques (40 – 45%)

Use the Concepts and Text Parsing Nodes to extract Terms and Concepts

- Use lists to include or exclude or combine terms (i.e. start, stop, synonym).
- Explain why Concepts are useful.
- Explain predefined Concepts.
- Define custom Concepts for a project.
- Modify the Term Table and explain the impact on the pipeline.
- Explain the impact of concepts on the pipeline.
- View document matches and similarity scores.
- Explore the term map (identify various components).

Write Concept Rules

- Given a scenario, use LITI to write a rule to achieve a goal (i.e. CATEGORY, CLASSIFIER, CONCEPT, C_CONCEPT, CONCEPT_RULE, NO_BREAK, PREDICATE_RULE, REGEX, etc.)
- Given a LITI rule, explain the how it influences scoring documents.
- Given a LITI rule, explain the how it impacts the term table.
- Identify and correct common syntax errors

Use the Topics Node to extract machine-generated topics

- Given a scenario, appropriately adjust term density.
- Given a scenario, appropriately adjust document density.
- Promote a topic to a category.
- Split and merge topics.
- Edit Topic Properties.
- Create Custom Topics.

Use rules to identify documents belonging to specific categories

- Analyze categorization results (F-Measure, precision, recall, misclassification).
- Edit and enhance predefined rules with defined concepts.
- Explain Categories Results (diagnostic counts, diagnostic metrics, categories score code).
- View Document matches and sentiment score.
- Explain sentiment level scoring.

Write Category rules

- Given a scenario, create and run appropriate Boolean rules to achieve a goal.
- Given a rule, explain how it impacts document categorization.

Use a Recurrent Neural Network (RNN) to recognize patterns

- Build a Basic RNN.
- Build a Bi-directional RNN.
- Build a Specialized (GRU, LSTM) RNN.

Identifying Image Patterns Using Computer Vision Techniques (35 – 40%)

Use convolutional layers in a Convolutional Neural Network (CNN)

- Explain the use of kernel filters in a CNN.

- Explain and calculate feature maps in a CNN (i.e. size).
- Detail equivariance to translation.
- Define hyperparameters (width, height and stride).
- Detail number of weights.

Use padding in a Convolutional Neural Network

- Detail the impact of padding on the feature map size.
- Use padding to accommodate skip-layer connections.
- Given a scenario, use padding to accomplish a goal.

Use pooling in a Convolutional Neural Network

- Detail the impact of pooling on the invariance of the CNN.
- Define summary functions used in pooling layers.
- Explain the use of filters in a CNN.
- Given a scenario, determine if using pooling is appropriate.

Use fully connected layers in a Convolutional Neural Network

- Given specific action calls, define number of parameters (trainable, estimated, etc.).
- Order FC layers correctly in building a CNN.
- Define activation functions used in Fc layers.

Use output layers in a Convolutional Neural Network

- Specify activation function for the output layer.
- Define types of error functions.

Tune the Hyperparameters of a Convolutional Neural Network

- Tune a deep learning model using the Hyperband method.

Score new image data

- Use trained weights to score new image data.
- Explain the relevance of batch size in scoring new image data.

Explain the impact of various architectural designs

- Use residual connections.
- Use concatenation connections.
- Define requirements for skip layer connections.
- Use one-by-one convolutions.
- Use Spatial Exploration techniques.

- Define blocks.
- Use cardinality techniques in the network structure.

Use regularization techniques

- Explain batch normalization.
- Use batch normalization to improve model generalization and learning.
- Explain dropout.
- Use dropout to improve model generalization.

Note: All 18 main objectives will be tested on every exam. The expanded objectives are provided for additional explanation and define the entire domain that could be tested.