



**USC Viterbi**  
School of Engineering

# Emotion Recognition Using Hierarchical Decision Tree Approach

**Chi-Chun Lee, Emily Mower, Carlos Busso, Sungbok Lee, Shrikanth Narayanan**  
**Signal Analysis and Interpretation Laboratory (SAIL)**  
**Electrical Engineering Department**  
**University of Southern California**



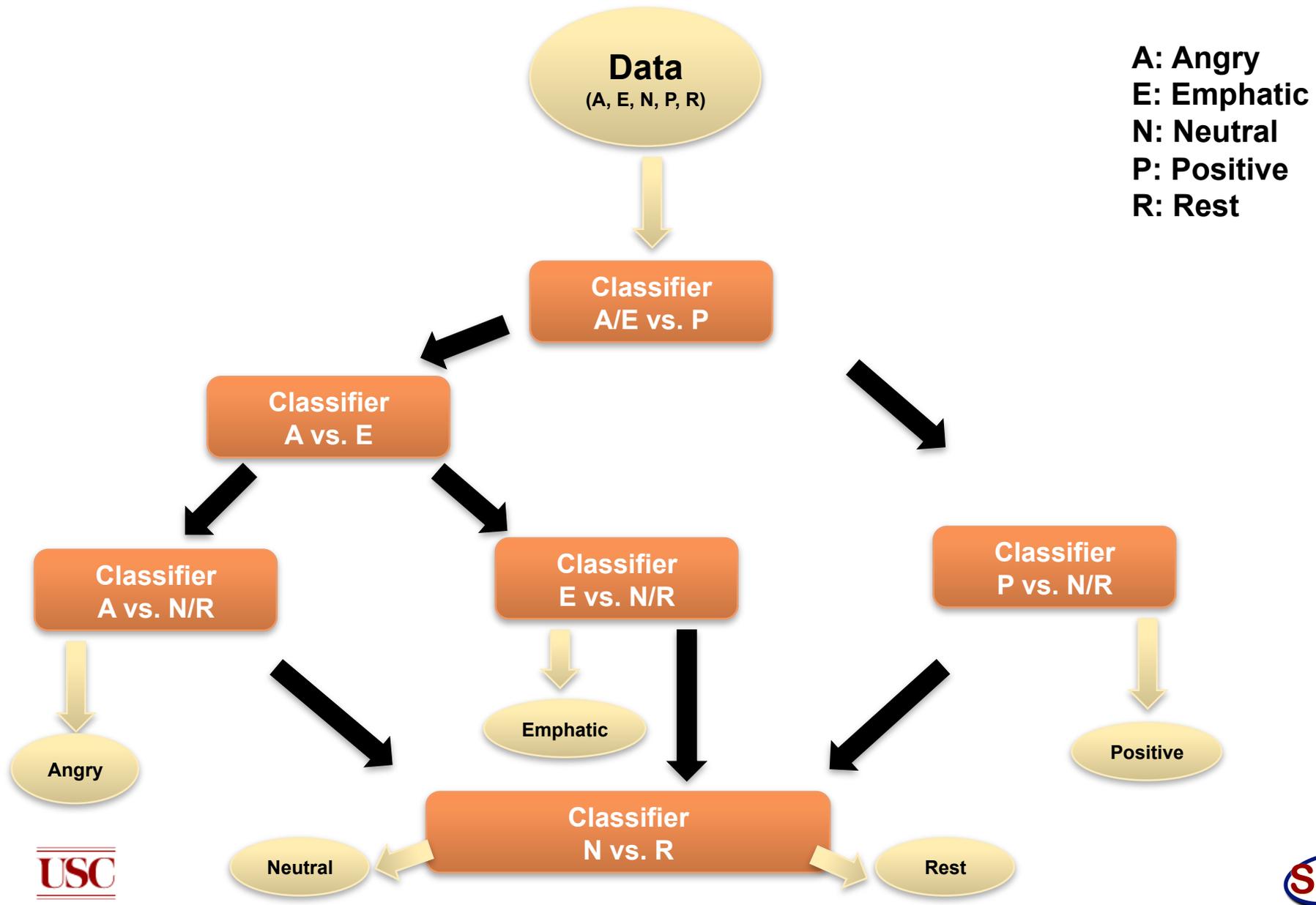
- **2009 Interspeech Emotion Challenge**
  - Classifier Sub-Challenge
  - Five Emotion Classes
    - Emphatic, Angry, Neutral, Positive, Rest
- **AIBO Database**
  - Training: 9959 – labeled
  - Testing: 8257 – unlabeled
  - 384 Dimension Feature Vector
- **Primary Performance Measure**
  - Unweighted Recall Percentage
  - Weighted Recall Percentage

# Proposed Method

**We used a combination of binary classifiers instead of one multi-class classifier**

**We proposed a classification framework composed of a hierarchical tree, where the top level classification was performed on the *easiest* emotion recognition task**

# Classification Framework



- **Bayesian Logistic Regression**
  - Binary Classifier
  - Gaussian prior on model coefficients  $\beta$
  - Threshold tuning based on balanced error rate criterion
- **Support Vector Machine**
  - Binary Classifier
  - Linear Kernel
  - Cost of Error,  $C$ , was set to approximate the distribution of data
- **Feature Selection**
  - Six sets of features
  - Forward feature selection using binary logistic regression
- **Feature Normalization**
  - *z-normalize* –  $\mu, \sigma^2$  calculated from the neutral data in the training dataset

# Experiment Setup



**Experiment I:** Leave one speaker out - 26 fold - cross validation on the training dataset

**Experiment II:** Evaluate performance on the unlabeled testing dataset

# Experiment I : Result & Discussion

## Bayesian Logistic Regression

Unweighted Recall

**48.27%**

Weighted Recall

48.82%

	Angry	Emphatic	Neutral	Positive	Rest
Angry	<b>504</b>	145	126	53	53
Emphatic	397	<b>1078</b>	412	101	107
Neutral	506	1020	<b>2703</b>	776	585
Positive	21	31	121	<b>439</b>	62
Rest	97	130	185	171	<b>138</b>

**94.82 % Recall**

# Experiment I : Result & Discussion

## Bayesian Logistic Regression

Unweighted Recall

**48.27%**

Weighted Recall

48.82%

	Angry	Emphatic	Neutral	Positive	Rest
Angry	<b>504</b>	145	126	53	53
Emphatic	397	<b>1078</b>	412	101	107
Neutral	506	1020	<b>2703</b>	776	585
Positive	21	31	121	<b>439</b>	62
Rest	97	130	185	171	<b>138</b>

- **Rest** is recognized at about chance level
- Comparable recall percentage for emotional classes (**Angry, Emphatic & Positive**)

# Experiment I : Result & Discussion

## Support Vector Machine

Unweighted Recall

**47.44%**

Weighted Recall

46.84%

	Angry	Emphatic	Neutral	Positive	Rest
Angry	<b>463</b>	159	123	57	79
Emphatic	322	<b>1041</b>	424	156	150
Neutral	386	930	<b>2548</b>	958	768
Positive	27	29	103	<b>446</b>	69
Rest	80	123	159	192	<b>167</b>

# Experiment II : Result & Discussion

## Submitted Results on Test Dataset

	Unweighted Recall	Weighted Recall
Baseline	38.20%	39.20%
Bayesian Logistic Regression	<b>41.57%</b>	<b>39.87%</b>
Support Vector Machine	40.84%	38.05%

	Angry	Emphatic	Neutral	Positive	Rest
Angry	<b>290</b>	171	65	63	22
Emphatic	210	<b>752</b>	325	136	150
Neutral	748	1094	<b>2047</b>	1109	369
Positive	23	13	39	<b>131</b>	9
Rest	95	58	134	197	<b>62</b>

- 3.37% absolute (8.82% relative) improvement (Bayesian Logistic Regression)
- 52% recall percentage for emotional classes (**Angry, Emphatic & Positive**)
- 25% recall percentage for non-emotional classes (**Neutral, Rest**)

- **Concluding Summary**
  - 3.37% absolute (8.82% relative) improvement over baseline
  - Carefully designed hierarchical structure
  - Easier classification task as first level
- **Future Work**
  - Other feature selection scheme (large-margin based)
  - Ensemble learning techniques
  - Soft decision at every decision level

# QUESTION & ANSWER