Ilya Kuzovkin

Adaptive Interactive Learning: a Novel Approach to Training Brain-Computer Interface Systems



Supervised by Konstantin Tretyakov

University of Tartu, 2013

Brain-Computer Interface



Mental intention



Brain-Computer Interface



Neuroimaging





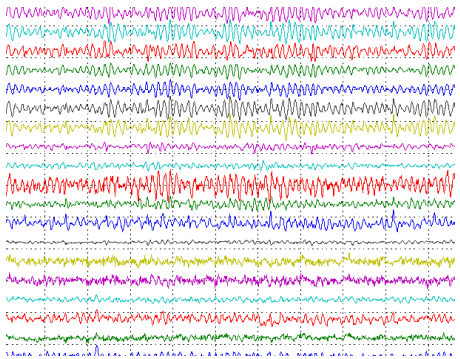
Mental intention



Brain-Computer Interface



Neuroimaging



Signal



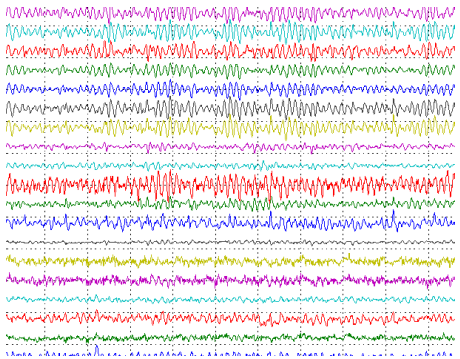
Mental intention



Brain-Computer Interface



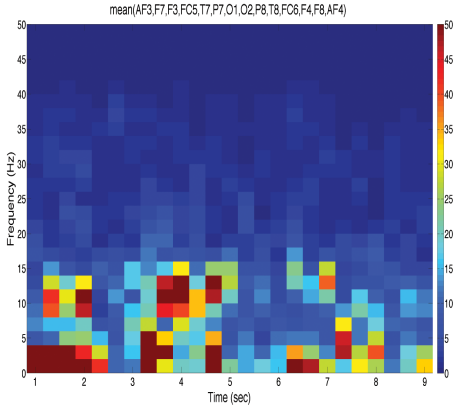
Neuroimaging



Signal



Mental intention



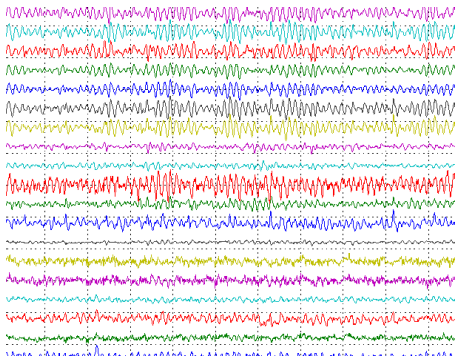
Representation



Brain-Computer Interface



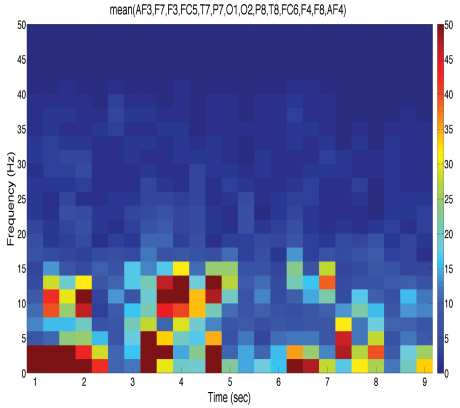
Neuroimaging



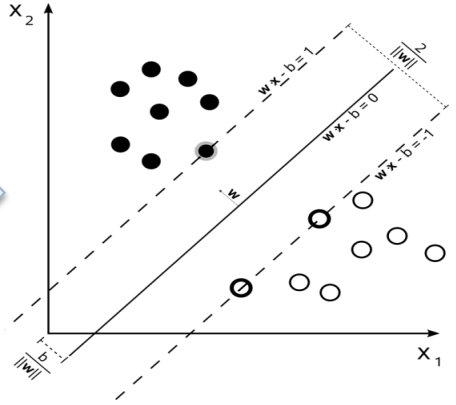
Signal



Mental intention



Representation

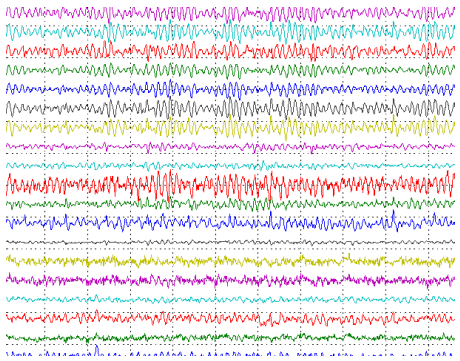


Algorithm

Brain-Computer Interface



Neuroimaging

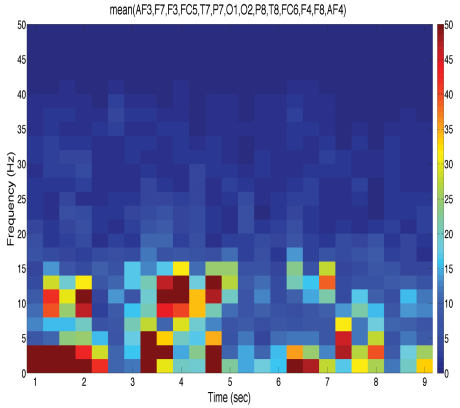


Signal

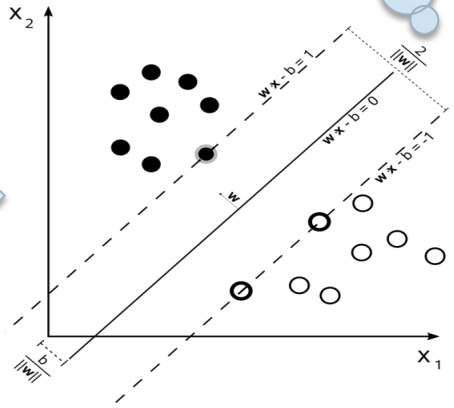
With 87% certainty I can say that you are thinking “Left” right now



Mental intention

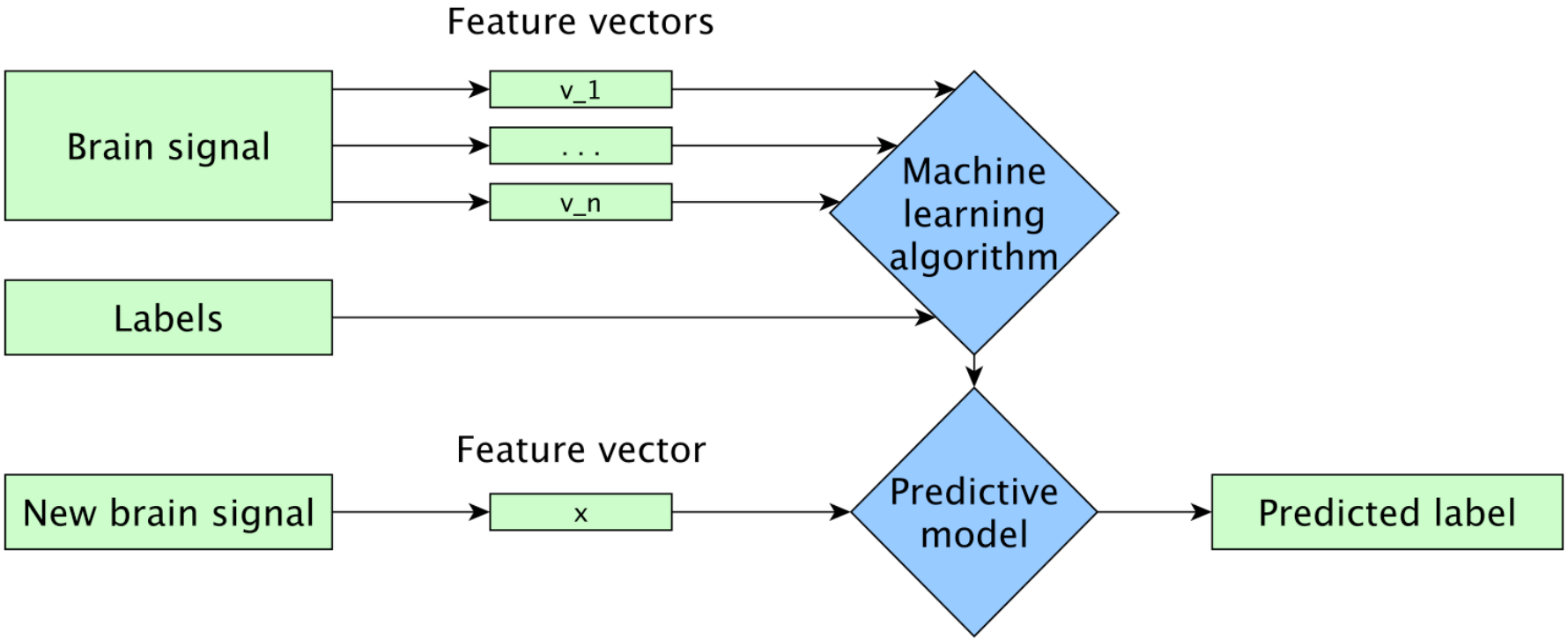


Representation



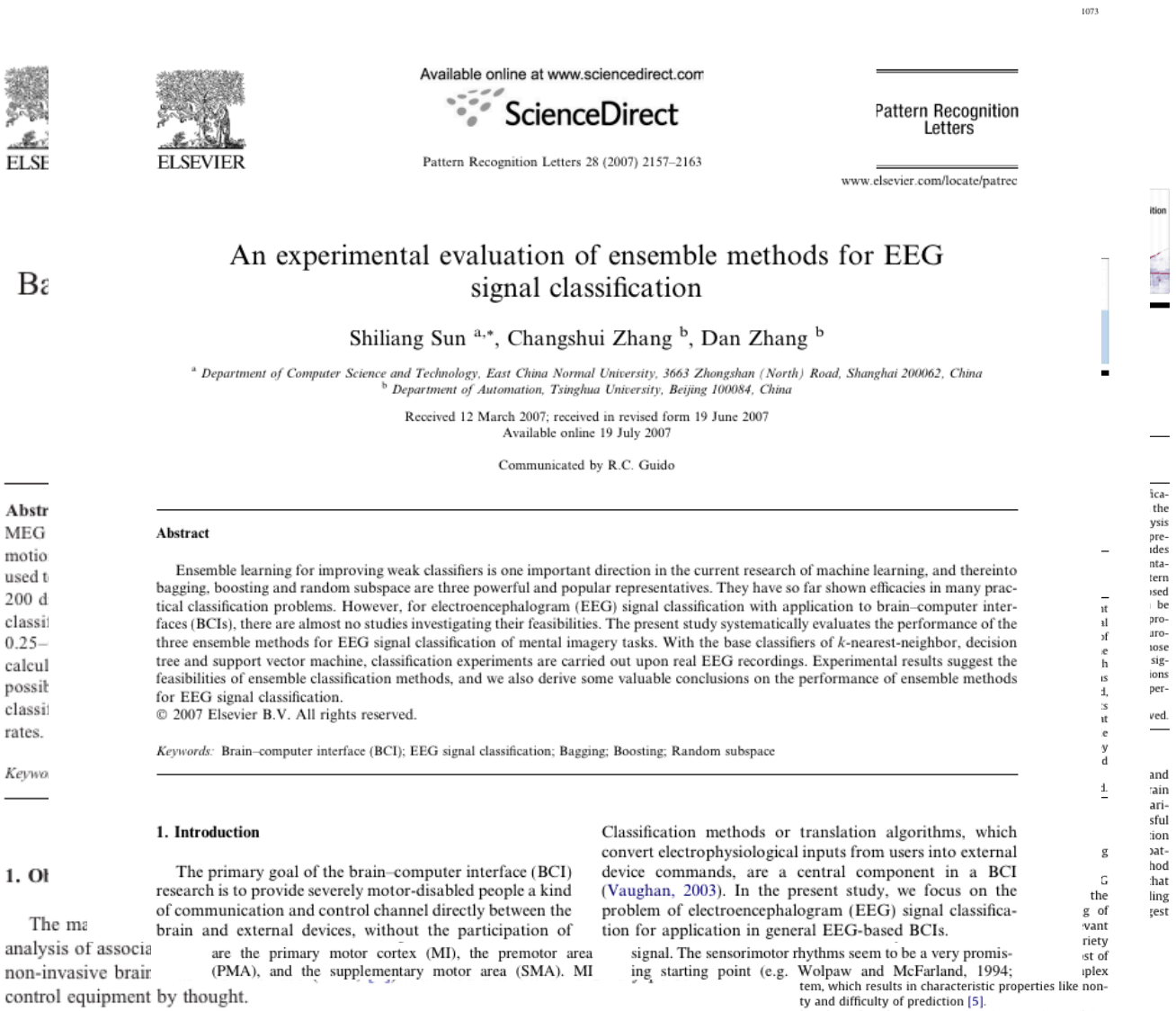
Algorithm

Traditional Approach



- User has to **pre-decide** which thoughts he will use to control the machine
- **No feedback** during the learning process

Two Problems



Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 28 (2007) 2157–2163
Pattern Recognition Letters
www.elsevier.com/locate/patrec

An experimental evaluation of ensemble methods for EEG signal classification

Shiliang Sun [a,*], Changshui Zhang [b], Dan Zhang [b]

[a] Department of Computer Science and Technology, East China Normal University, 3663 Zhongshan (North) Road, Shanghai 200062, China
[b] Department of Automation, Tsinghua University, Beijing 100084, China

Received 12 March 2007; received in revised form 19 June 2007
Available online 19 July 2007

Communicated by R.C. Guido

Abstract

Ensemble learning for improving weak classifiers is one important direction in the current research of machine learning, and thereinto bagging, boosting and random subspace are three powerful and popular representatives. They have so far shown efficacies in many practical classification problems. However, for electroencephalogram (EEG) signal classification with application to brain–computer interfaces (BCIs), there are almost no studies investigating their feasibilities. The present study systematically evaluates the performance of the three ensemble methods for EEG signal classification of mental imagery tasks. With the base classifiers of k-nearest-neighbor, decision tree and support vector machine, classification experiments are carried out upon real EEG recordings. Experimental results suggest the feasibilities of ensemble classification methods, and we also derive some valuable conclusions on the performance of ensemble methods for EEG signal classification.
© 2007 Elsevier B.V. All rights reserved.

Keywords: Brain–computer interface (BCI); EEG signal classification; Bagging; Boosting; Random subspace

1. Introduction

The primary goal of the brain–computer interface (BCI) research is to provide severely motor-disabled people a kind of communication and control channel directly between the brain and external devices, without the participation of

are the primary motor cortex (MI), the premotor area (PMA), and the supplementary motor area (SMA). MI

Classification methods or translation algorithms, which convert electrophysiological inputs from users into external device commands, are a central component in a BCI (Vaughan, 2003). In the present study, we focus on the problem of electroencephalogram (EEG) signal classification for application in general EEG-based BCIs.

signal. The sensorimotor rhythms seem to be a very promising starting point (e.g. Wolpaw and McFarland, 1994;

tem, which results in characteristic properties like nonty and difficulty of prediction [5].

2-class accuracy

≈ 0.9

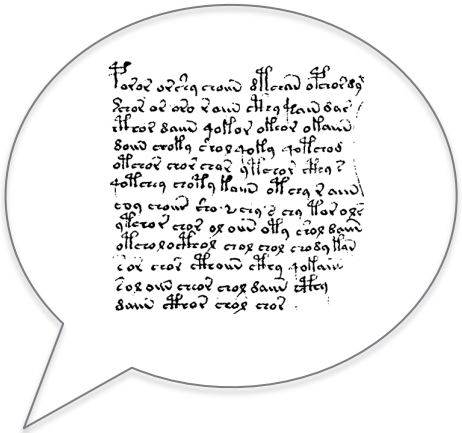
3-class accuracy

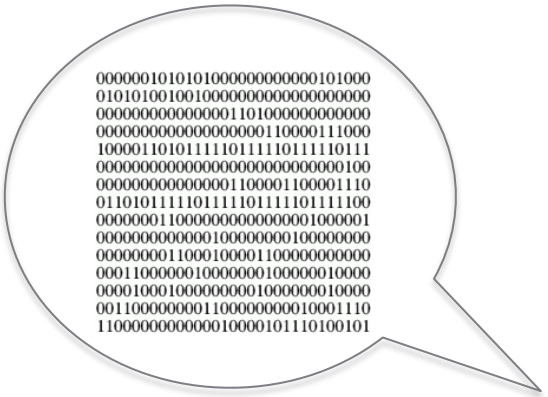
≈ 0.7

Inconsistency
of the signal

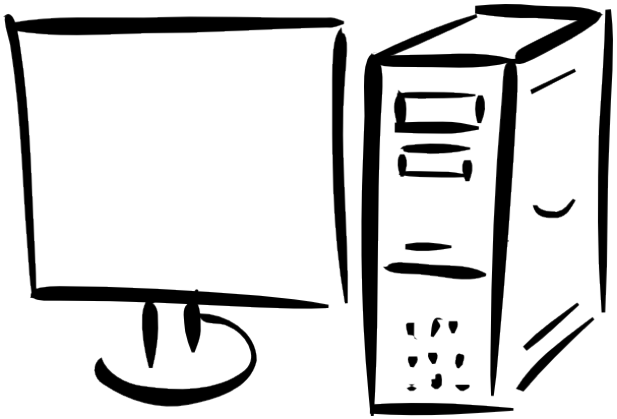
Producing
Distinguishable
mental states

Abstract Concept: Communication

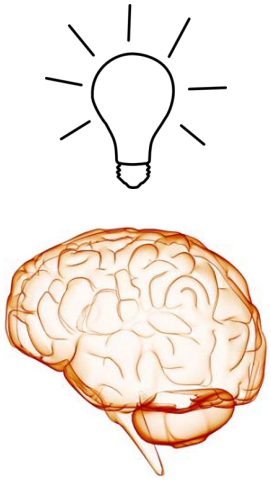


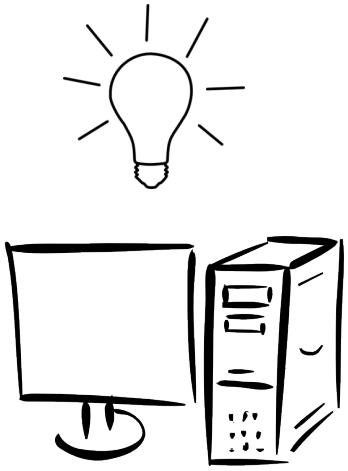






Establishing communication protocol





000000101010100000000000101000

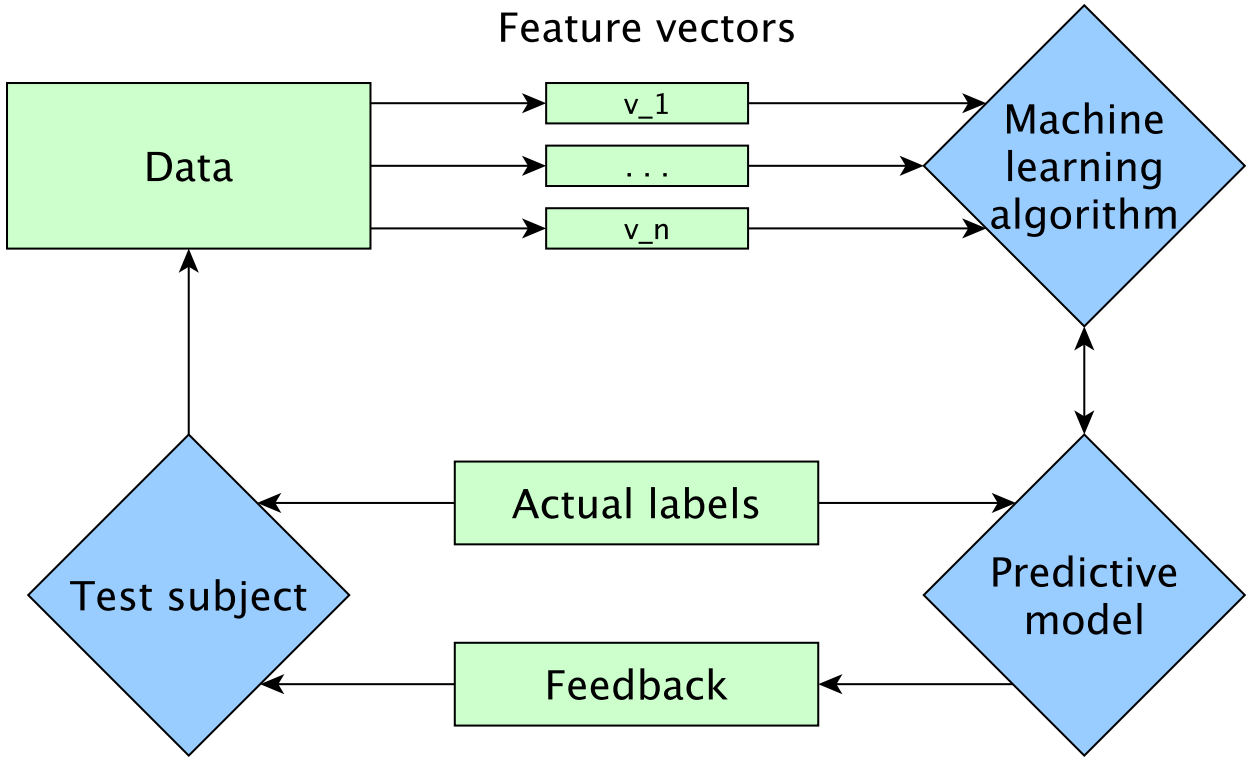
100001101011111011111011111110111

You want to move “Left”?

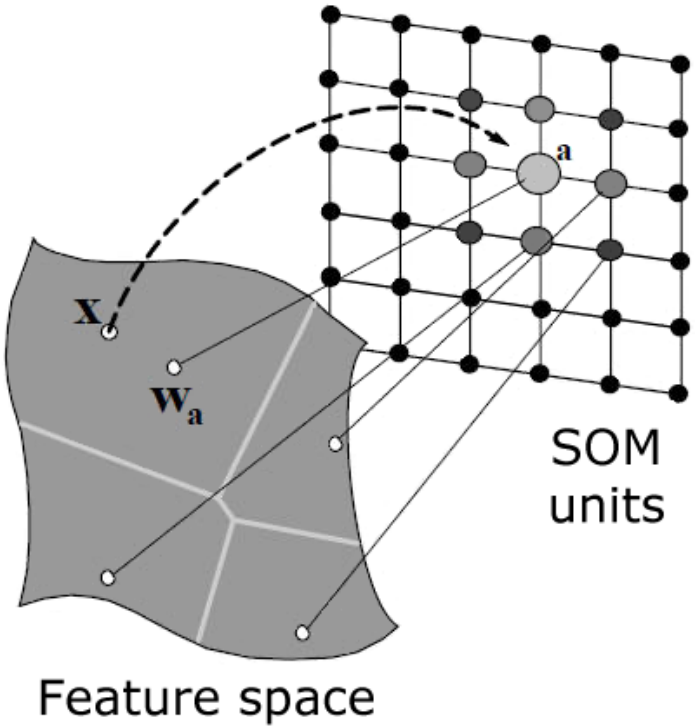
Yes!

Adaptive Interactive Learning

- Data is processed in **real-time**
- After **each** sample both the user and the machine get **feedback**
- User can **adapt** his behavior according to feedback
- Machine can **update** the predictive model



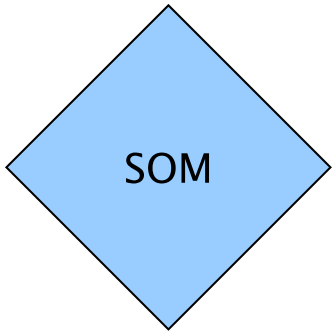
Self-Organizing Map



- **Unsupervised** learning algorithm
- Map consists of **units**
- Weight vector **w**(a) maps units to feature space
- For a new data vector **x** we find the **best matching unit**: the one, which has weight vector **w**(a) closest to **x**

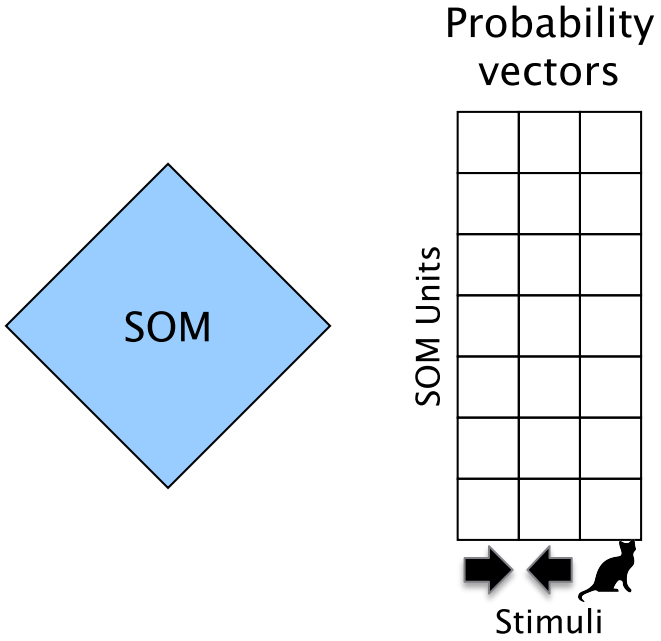
Predictive SOM

- Class probability vector $\mathbf{p}(a) = (\Pr[\text{action}_1|a], \ldots, \Pr[\text{action}_n|a])$
- Probabilistic confusion matrix
- F1 score provides the feedback for the machine
- Feedback to the user is based on the class probabilities



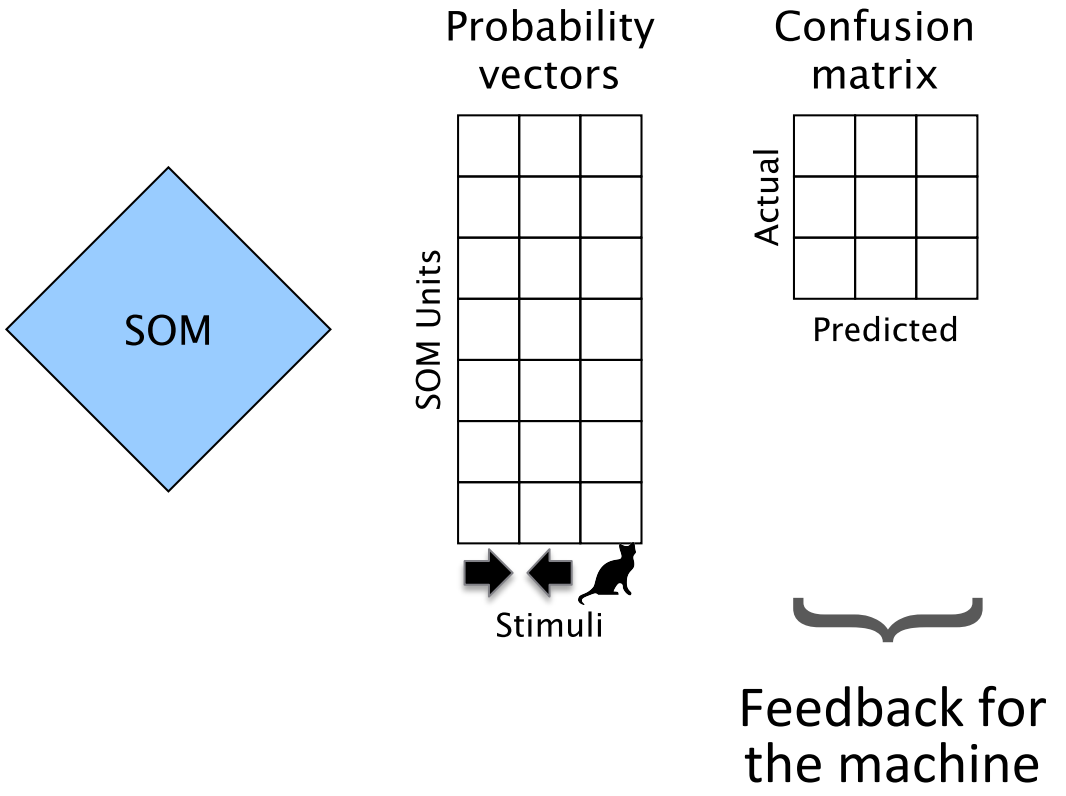
Predictive SOM

- Class probability vector $\mathbf{p}(a) = (\Pr[\text{action}_1|a], \ldots, \Pr[\text{action}_n|a])$
- Probabilistic confusion matrix
- F1 score provides the feedback for the machine
- Feedback to the user is based on the class probabilities



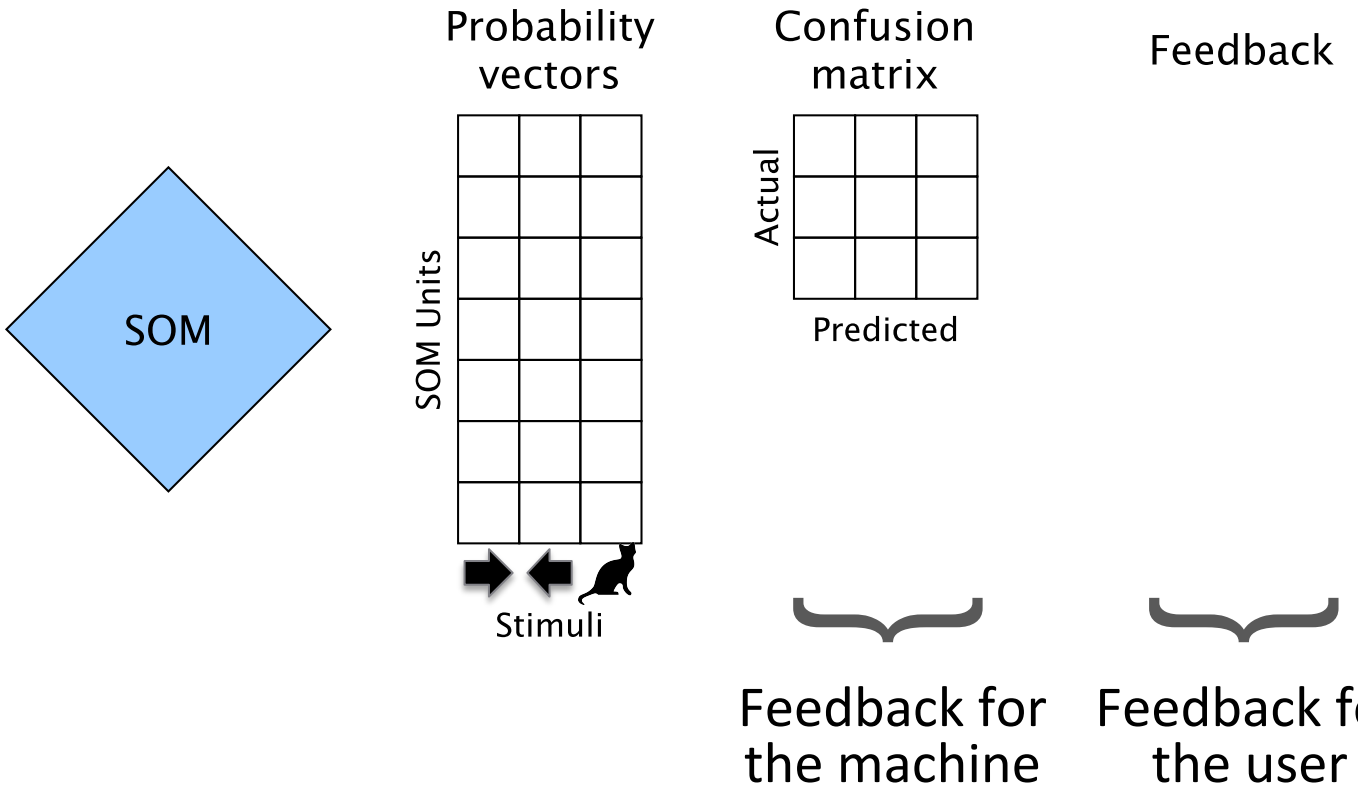
Predictive SOM

- Class probability vector $\mathbf{p}(a) = (\Pr[\text{action}_1|a], \ldots, \Pr[\text{action}_n|a])$
- Probabilistic confusion matrix
- F1 score provides the feedback for the machine
- Feedback to the user is based on the class probabilities



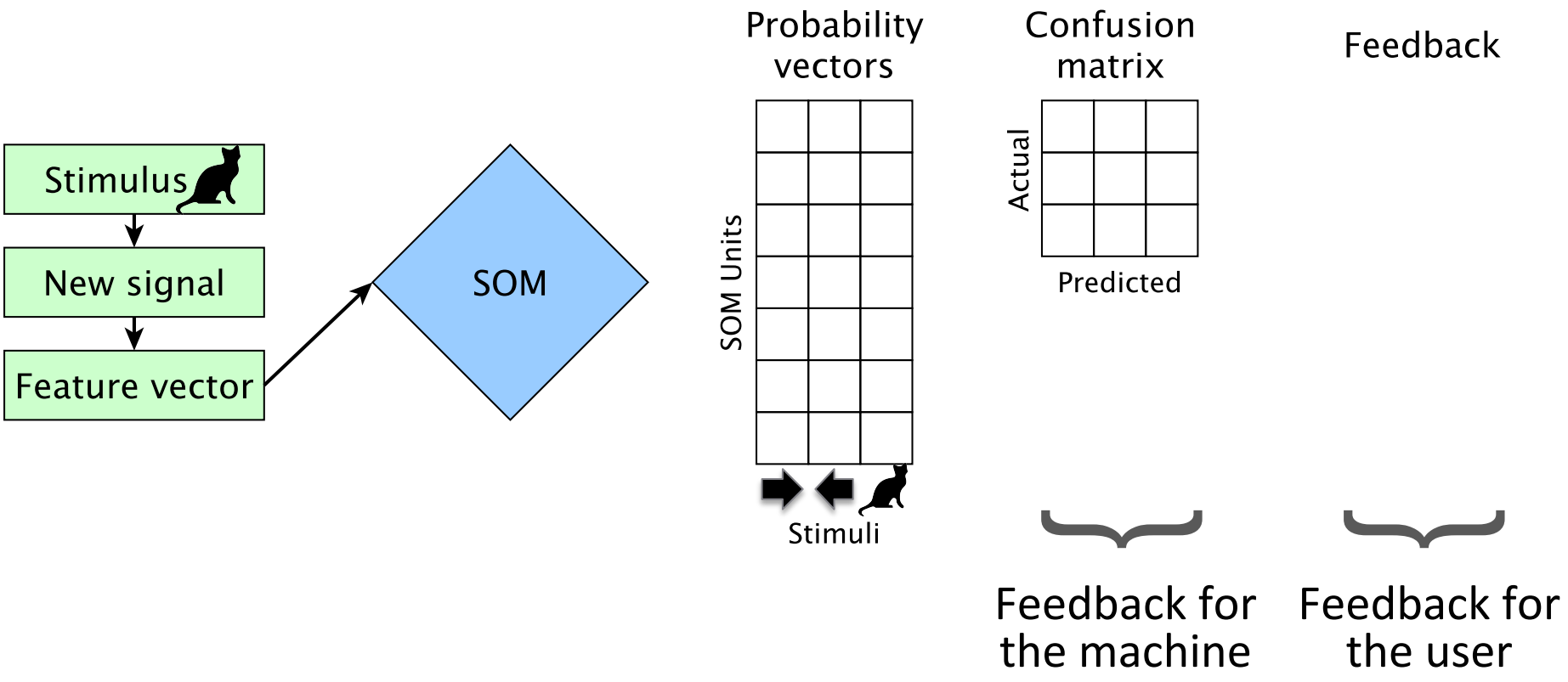
Predictive SOM

- Class probability vector $\mathbf{p}(a) = (\Pr[\text{action}_1|a], \ldots, \Pr[\text{action}_n|a])$
- Probabilistic confusion matrix
- F1 score provides the feedback for the machine
- Feedback to the user is based on the class probabilities



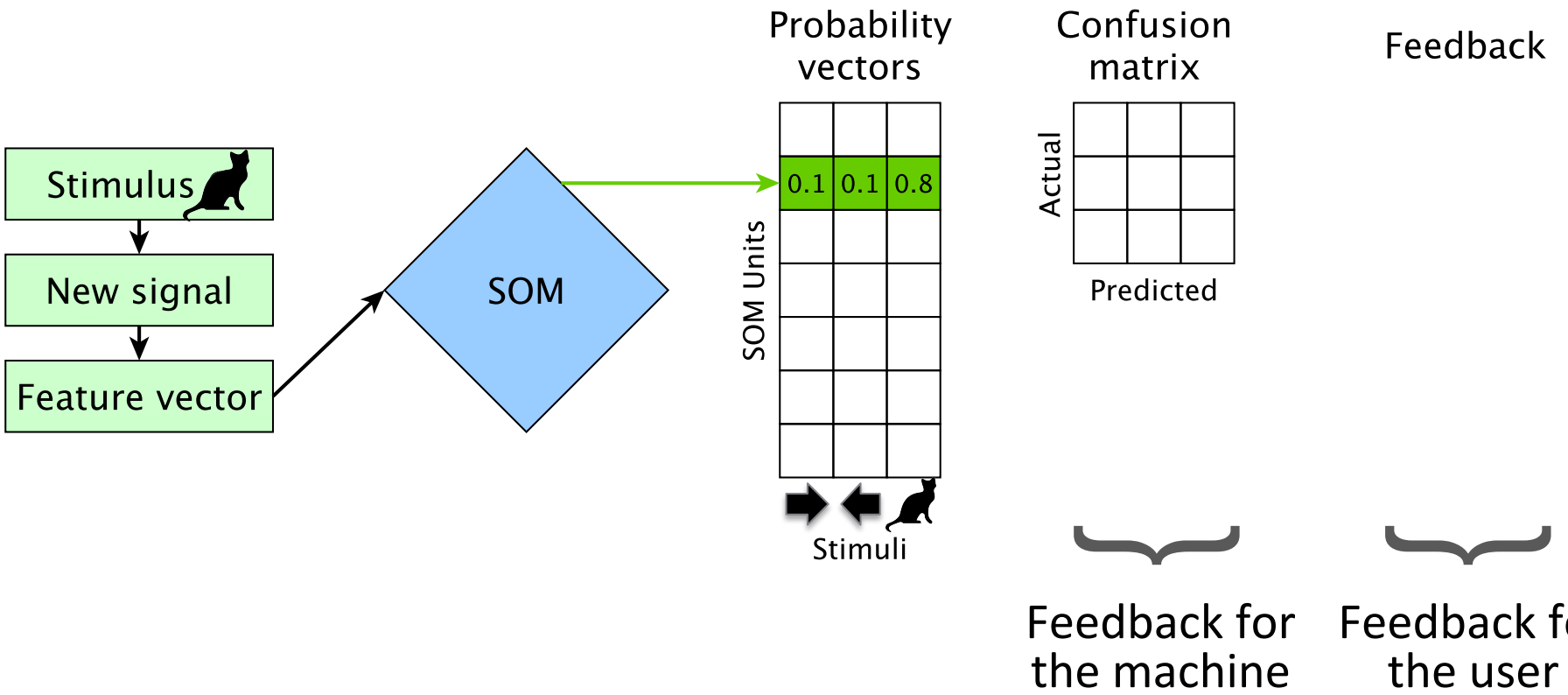
Predictive SOM

- Class probability vector $\mathbf{p}(a) = (\Pr[\text{action}_1|a], \ldots, \Pr[\text{action}_n|a])$
- Probabilistic confusion matrix
- F1 score provides the feedback for the machine
- Feedback to the user is based on the class probabilities



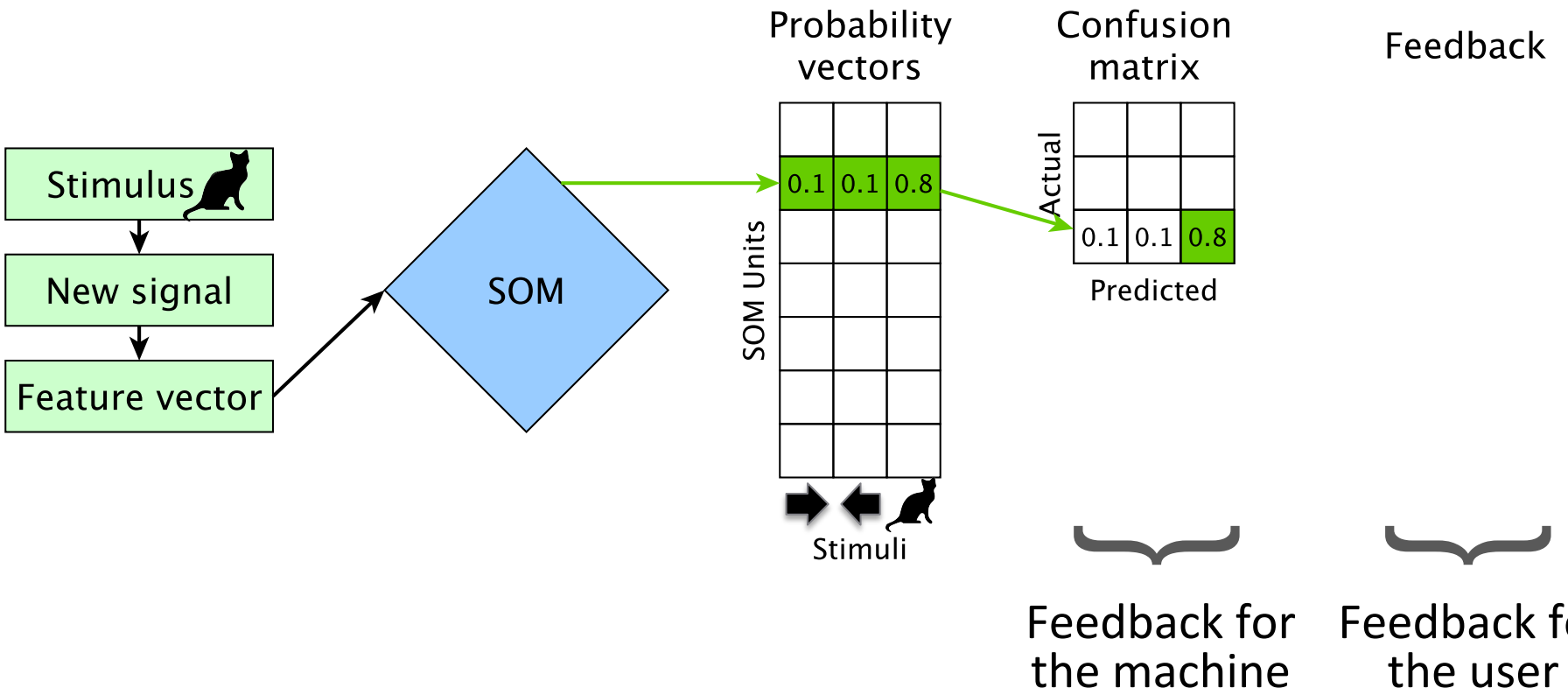
Predictive SOM

- Class probability vector $\mathbf{p}(a) = (\Pr[\text{action}_1|a], \ldots, \Pr[\text{action}_n|a])$
- Probabilistic confusion matrix
- F1 score provides the feedback for the machine
- Feedback to the user is based on the class probabilities



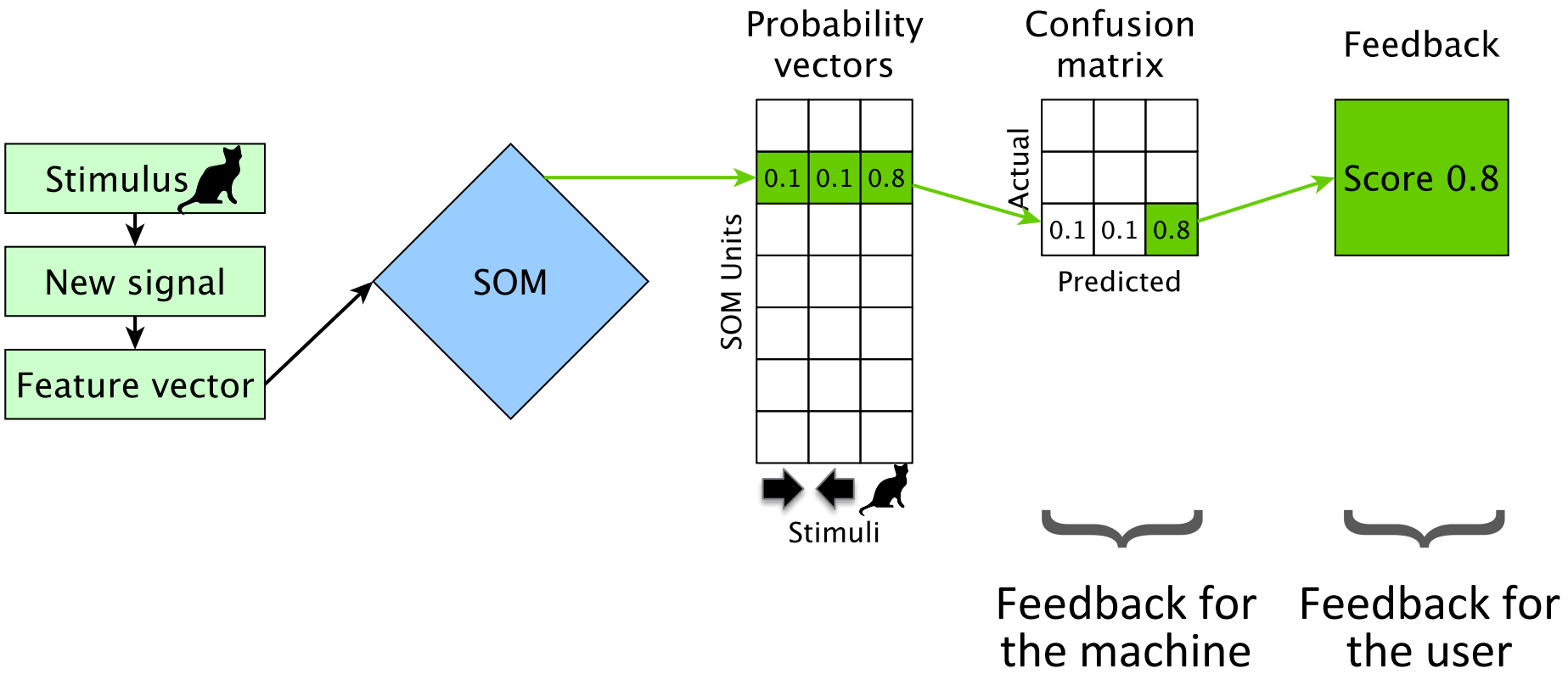
Predictive SOM

- Class probability vector $\mathbf{p}(a) = (\Pr[\text{action}_1|a], \ldots, \Pr[\text{action}_n|a])$
- Probabilistic confusion matrix
- F1 score provides the feedback for the machine
- Feedback to the user is based on the class probabilities



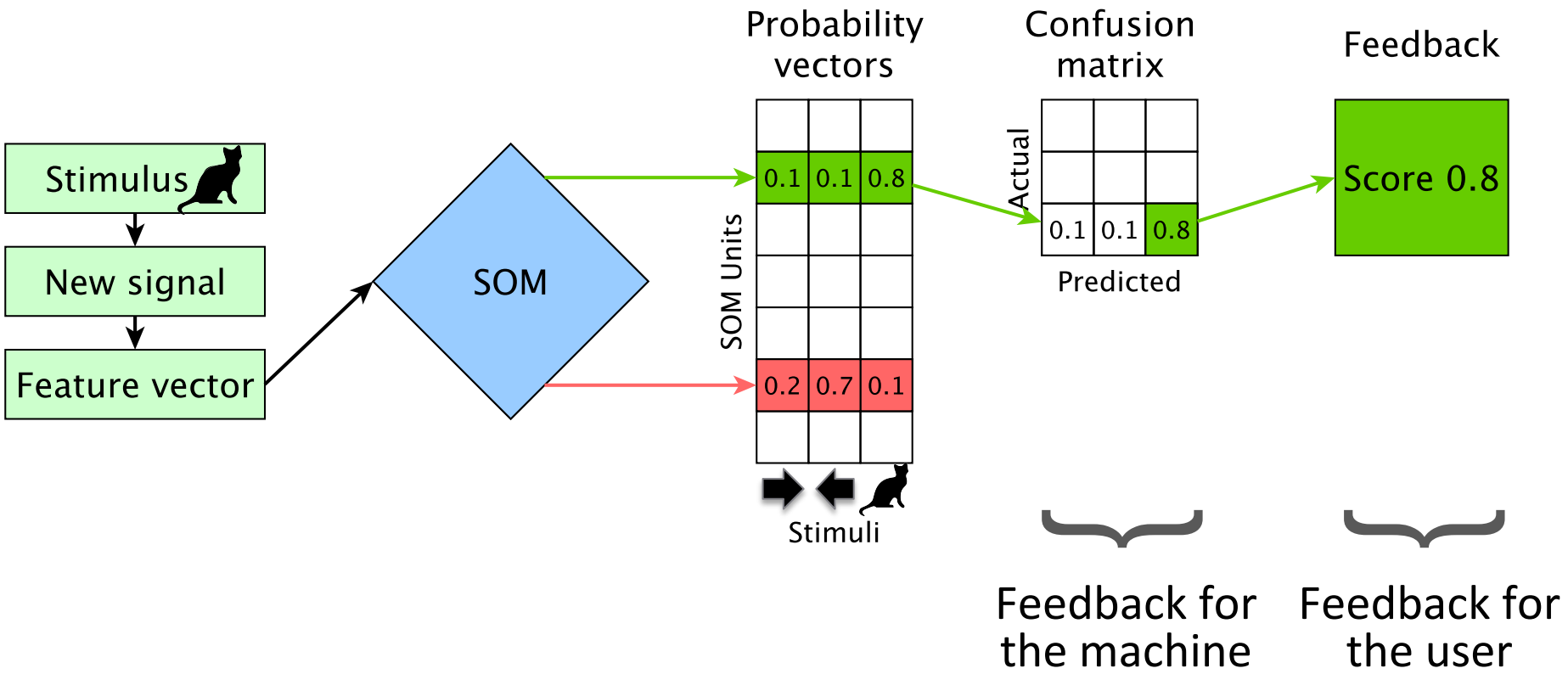
Predictive SOM

- Class probability vector $\mathbf{p}(a) = (\Pr[\text{action}_1|a], \ldots, \Pr[\text{action}_n|a])$
- Probabilistic confusion matrix
- F1 score provides the feedback for the machine
- Feedback to the user is based on the class probabilities



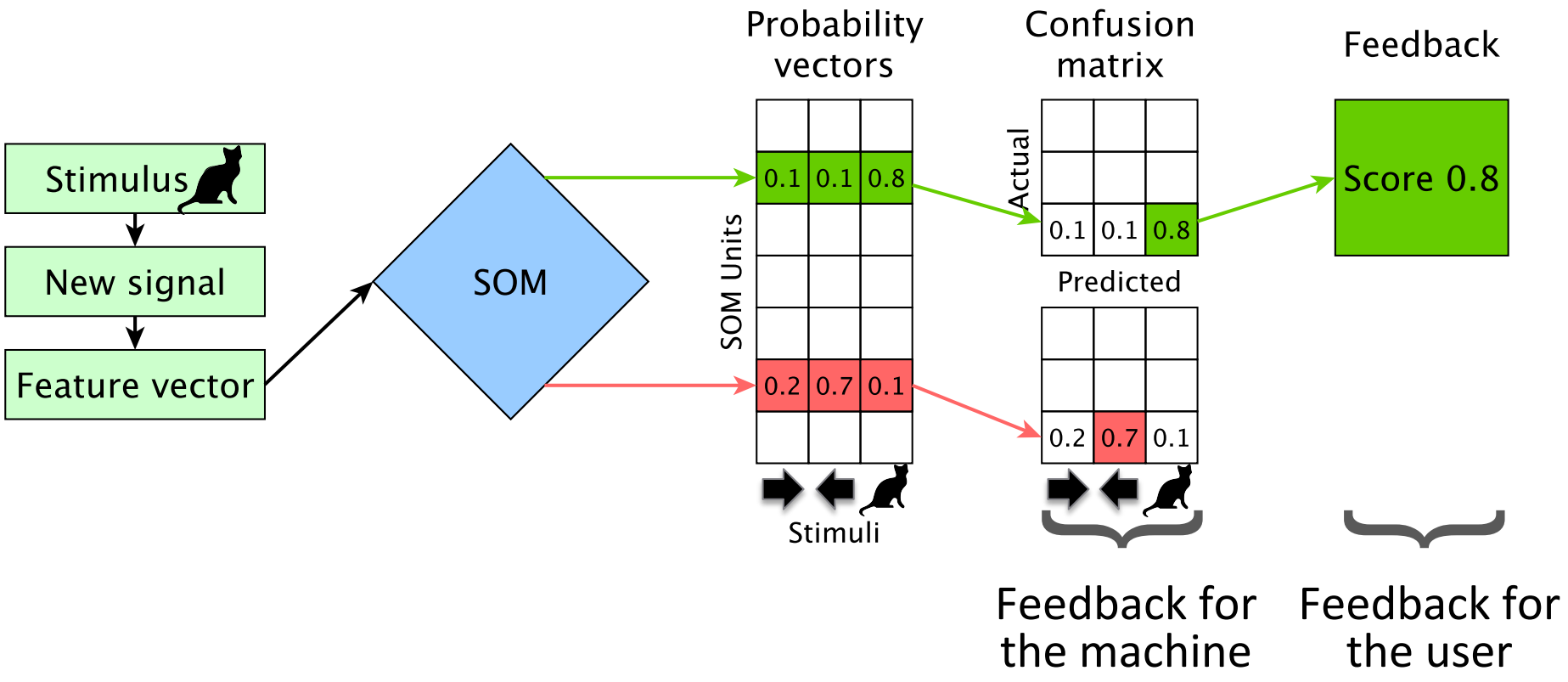
Predictive SOM

- Class probability vector $\mathbf{p}(a) = (\Pr[\text{action}_1|a], \ldots, \Pr[\text{action}_n|a])$
- Probabilistic confusion matrix
- F1 score provides the feedback for the machine
- Feedback to the user is based on the class probabilities



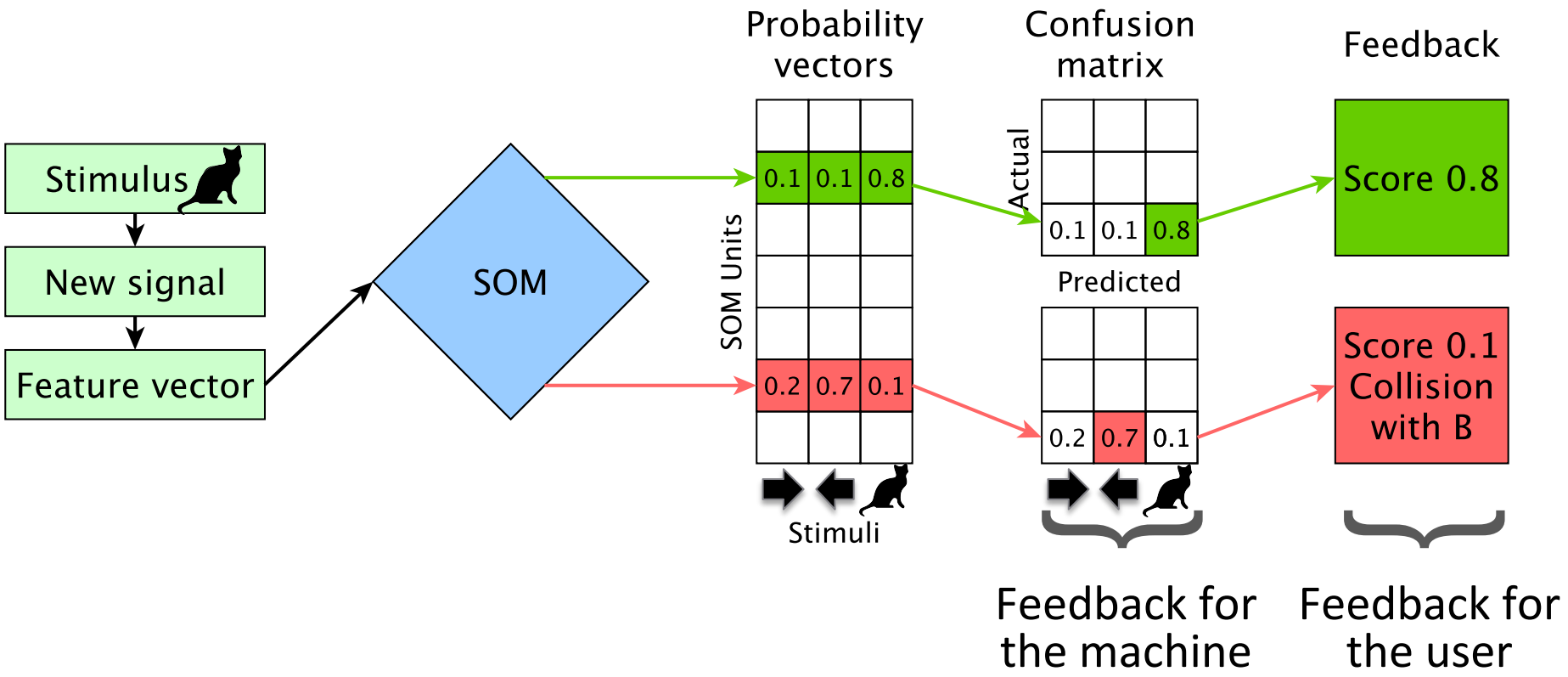
Predictive SOM

- Class probability vector $\mathbf{p}(a) = (\Pr[\text{action}_1|a], \ldots, \Pr[\text{action}_n|a])$
- Probabilistic confusion matrix
- F1 score provides the feedback for the machine
- Feedback to the user is based on the class probabilities

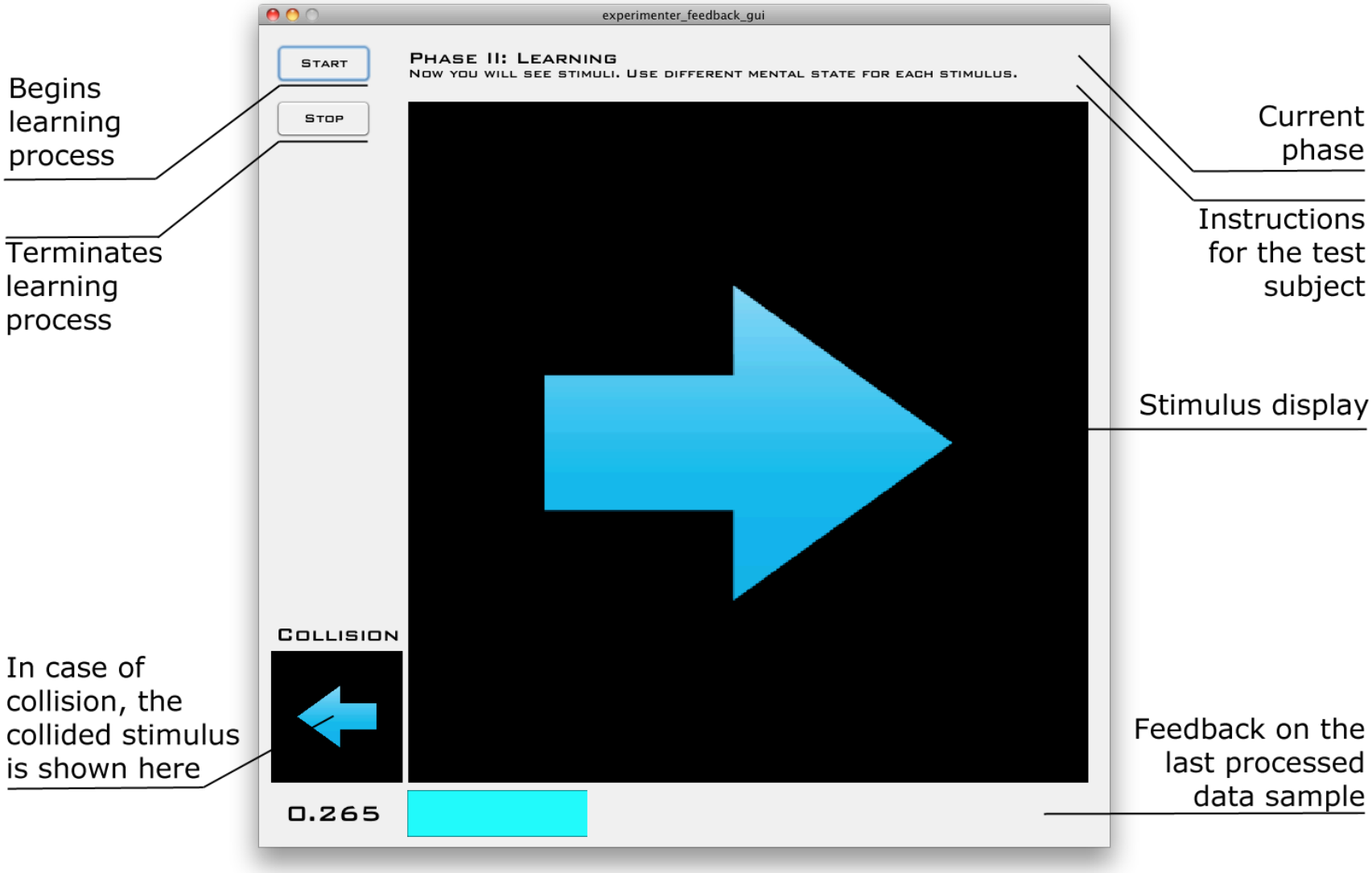


Predictive SOM

- Class probability vector $\mathbf{p}(a) = (\Pr[\text{action}_1|a], \ldots, \Pr[\text{action}_n|a])$
- Probabilistic confusion matrix
- F1 score provides the feedback for the machine
- Feedback to the user is based on the class probabilities

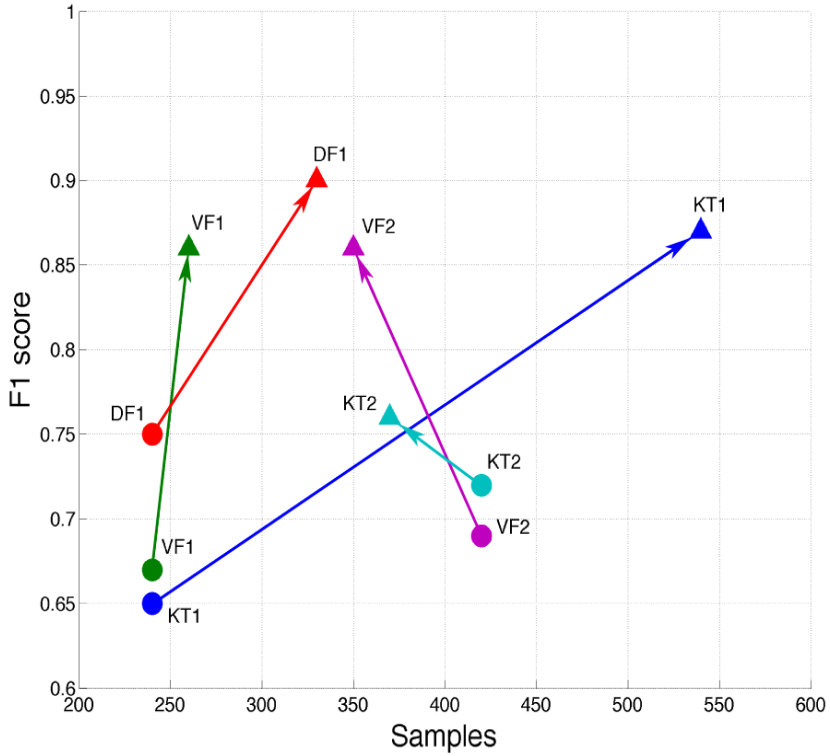


The Experimenter



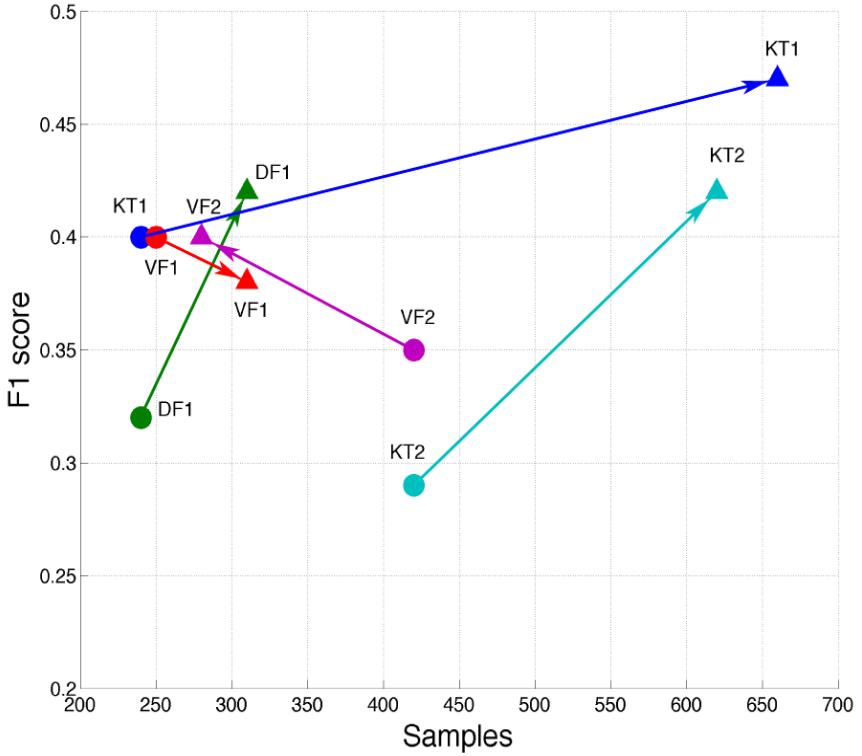
Traditional vs. Adaptive: Real Data

Facial expressions



Traditional	0.696
Adaptive	0.85
p-value	0.0075

Mental states



Traditional	0.352
Adaptive	0.418
p-value	0.06

Summary

- Proposed a **new** approach to BCI training
- **Implemented** the idea using Predictive SOM: **unsupervised online** learning, **one parameter**
- **Application** which embodies the new approach
- Demonstrated the **advantage** of the adaptive approach on **artificial** data
- Experimental results:
  - **0.07** F1 score increase for **mental states**
  - **0.15** F1 score increase for **facial expressions**