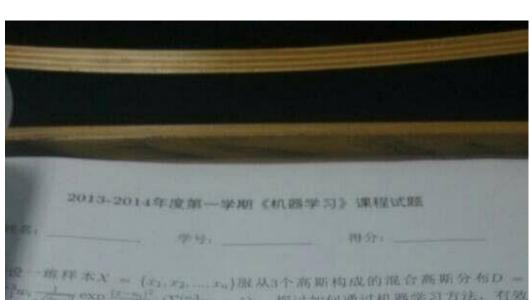
## 2012-2013年度第一学期《机器学习》课程试题

姓名:	学号:	得分:
XT-YET	3 3 .	

- 1. 假设一维样本 $X = \{x_1, x_2, ..., x_n\}$ 服从一维正态分布 $N(\mu, \sigma^2)$ ,试描述用最大似然估计的方法估计高斯分布的均值和方差的过程,并指出其局限性。(15分)
- 2. 试阐述Fisher's Linear Discriminant Analysis (LDA) 的基本原理,如何将LDA应用于多类问题。(25分)
- 3. 支持向量机 (Support Vector Machine) 是一种热门的机器学习算法, Adaboost 也是一种热门的机器学习算法, 如何将这两个方法应用于人脸检测, 并探讨 这两种方法的区别以及在性能上可能出现的差异。(25分)
- 4. 试以K均值聚类为例, 描述期望最大化算法(EM算法)的基本原理。(15分)
- 5. 随着信息化的发展,大数据的时代已经到来。海量的文本、图像、视频数据 存在于互联网上,请结合自己的科研背景和兴趣,探讨机器学习方法如何在 大数据分析、处理中应用。(20分)



 $\frac{1}{\sqrt{2\pi\sigma^2}} \exp{\frac{(x-a\mu)^2}{2\sigma^2}}$   $(\Sigma_{i=1}^{(a)}w_i=1)$  。探讨如何通过机器学习方法。有效 这三个高斯分布的参数和权重。(20分)

份分析 (Principal Component Analysis) 是一种典型的线性降维方法 含一种具体应用探讨主成份分析的思想, 并探讨其局限性和改进 2015/1

量机 (Support Vector Machine) 是一种热门的机器学习算法。专 题为例描述支持向量机 (SVM) 算法的基本原理, 并探讨如何 5类问题。(20分)

iscrete AdaBoost的基本原理,并尝试探讨如何利用Discret 亍垃圾邮件过滤。(20分)

己的科研兴趣和方向, 谈谈能否有合适的机器学习算法在 来哪些改进,存在哪些难点。(20分)