EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Mathematisch-Naturwissenschaftliche Fakultät

Fachbereich Mathematik

Prof. Dr. Andreas Prohl
Armin Beck

Statistical Learning 1

Summer semester 2024 **Tübingen, 03.06.2024**

Assignment 7

Problem 1
The third regression estimator for which we prove weak universal consistency is the k-NN estimator, provided that

$$k_n \uparrow \infty \text{ and } \frac{k_n}{n} \downarrow 0 \text{ for } n \uparrow \infty$$

Verify assumptions (ii) and $(iv)_2$ in Stone's theorem.

Problem 2
We call an observation noiseless if $Y_i = m(X_i)$, for $1 \leq i \leq n$. Prove that for fixed k the k-NN regression estimator is weakly universal consistent for noiseless observations.

Problem 3
Fix $x \in \mathbb{R}^d$. Let g_n be the k-NN classification rule for M-classes:

$$g_n(x) = \text{argmax}_{1 \leq j \leq M} \sum_{i=1}^{k_n} \mathbb{1}_{\{y_{(n,i}(x)=j\}}.$$

Show that, for $k_n \uparrow \infty$ and $\frac{k_n}{n} \downarrow 0$,

$$\lim_{n\to\infty} \mathcal{P}[\{g_n(X) \neq Y\}] = \mathcal{P}[\{g^*(X) \neq Y\}]$$

for all distributions of (X, Y), where g^* is the Bayes decision rule.
Hint: Use **Problem 1 (b)** of **Assignment 2**, and the weak universal consistency property of the $k-$NN estimator.

Date of Submission: 10.06.2024 in the mailbox at 12 noon.