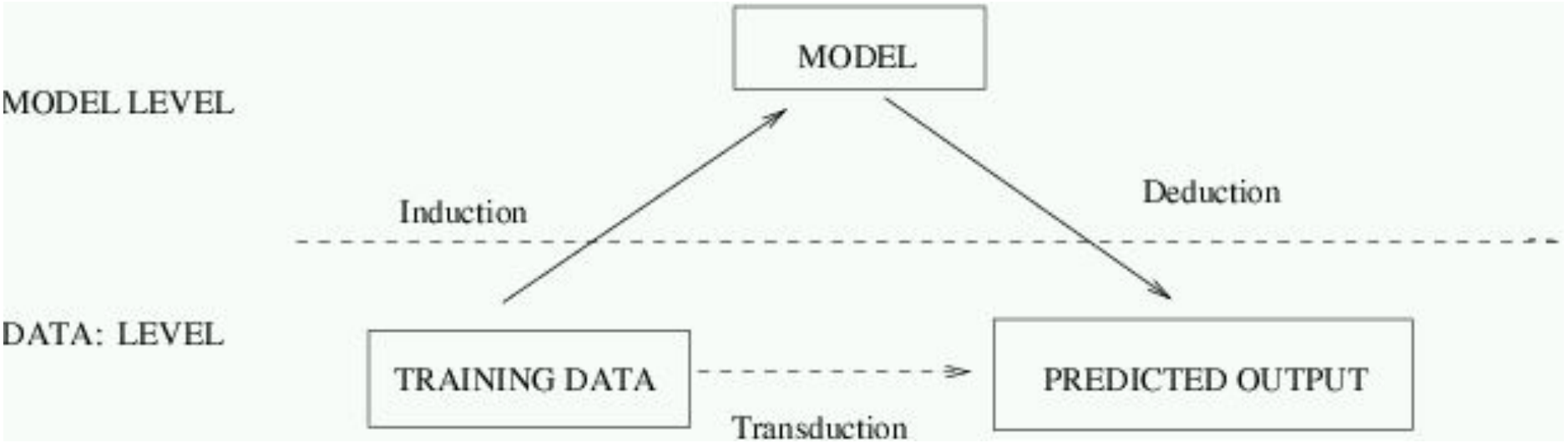
K-NN

- Today we will cover the KNN classifier and previous feature extraction methods.



http://www.cs.joensuu.fi/~whamalai/skc/ml.html

K-NN Classifier

- K Nearest Neighbor Classifier
- Aim is find the class label of a test sample
- As it clear from name, the samples are sorted with respect to their distance measures between test sample.
- Then the K samples are analyzed to make a decision.
- K must be an odd number as 1, 3, 5, 7, 9
- K=1. means the closest sample is the class label of processed test sample.

K-NN Classifier

- K=3, closest 3 samples are analyzed. Mode of samples indicates the class label of processed test sample.

- K=5, closest 5 samples are analyzed. Mode of samples indicates the class label of processed test sample.

K-NN Example

Samples	Feature1 (F1)	Feature2(F2)	Class Name
S1	100	5	Class1
S2	95	10	Class1
S3	78	110	Class1
S4	10	90	Class1
S5	77	85	Class1
S6	50	45	Class2
S7	23	100	Class2
S8	30	30	Class2
S9	40	60	Class2
S10	20	70	Class2
S11	5	0	Class3
S12	100	4	Class3
S13	1	20	Class3
S14	1	1	Class3
S15	22	71	Class3
Test	75	50	?

K-NN Classifier

- Assume distance metric is L2 distance

$$d = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \cdots + (q_n - p_n)^2}$$

- For this study

$$d = \sqrt{(\text{Test}_{F1} - \text{Train}_{F1})^2 + (\text{Test}_{F2} - \text{Train}_{F2})^2}$$

K-NN Example

Samples	Distance	Class Name
S1	51.48	1
S2	44.72	1
S3	60.07	1
S4	76.32	1
S5	35.06	1
S6	25.50	2
S7	72.14	2
S8	49.24	2
S9	36.40	2
S10	58.52	2
S11	86.02	3
S12	52.35	3
S13	79.85	3
S14	88.75	3
S15	57.01	3
Test	X= 75,50	?

Distances before sorted

K-NN Example

Samples	Distance	Class Name
S1	25.50	2
S2	35.06	1
S3	36.40	2
S4	44.72	1
S5	49.24	2
S6	51.48	1
S7	52.35	3
S8	57.01	3
S9	58.52	2
S10	60.07	1
S11	72.14	2
S12	76.32	1
S13	79.85	3
S14	86.02	3
S15	88.75	3
Test	75,50	?

Distances after sorted

K-NN Example

Samples	Distance	Class Name
S1	25.50	2
S2	35.06	1
S3	36.40	2
S4	44.72	1
S5	49.24	2
S6	51.48	1
S7	52.35	3
S8	57.01	3
S9	58.52	2
S10	60.07	1
S11	72.14	2
S12	76.32	1
S13	79.85	3
S14	86.02	3
S15	88.75	3
Test	75	?

K=1 then test sample assigned to class2

K-NN Example

Samples	Distance	Class Name
S1	25.50	2
S2	35.06	1
S3	36.40	2
S4	44.72	1
S5	49.24	2
S6	51.48	1
S7	52.35	3
S8	57.01	3
S9	58.52	2
S10	60.07	1
S11	72.14	2
S12	76.32	1
S13	79.85	3
S14	86.02	3
S15	88.75	3
Test	75	?

K=3 then test sample assigned to class2

Since mode is 2 for class2

K-NN Example

Samples	Distance	Class Name
S1	25.50	2
S2	35.06	1
S3	36.40	2
S4	44.72	1
S5	49.24	2
S6	51.48	1
S7	52.35	3
S8	57.01	3
S9	58.52	2
S10	60.07	1
S11	72.14	2
S12	76.32	1
S13	79.85	3
S14	86.02	3
S15	88.75	3
Test	75	?

K=5 then test sample assigned to class2

Since mode is 3 for class2

K-NN Example

Samples	Distance	Class Name
S1	25.50	2
S2	35.06	1
S3	36.40	2
S4	44.72	1
S5	49.24	2
S6	51.48	1
S7	52.35	3
S8	57.01	3
S9	58.52	2
S10	60.07	1
S11	72.14	2
S12	76.32	1
S13	79.85	3
S14	86.02	3
S15	88.75	3
Test	75	?

K=7 then test sample assigned to class1 or to class2
Since mode is 3 for class2 or class1

K-NN Example

Samples	Distance	Class Name
S1	25.50	2
S2	35.06	1
S3	36.40	2
S4	44.72	1
S5	49.24	2
S6	51.48	1
S7	52.35	3
S8	57.01	3
S9	58.52	2
S10	60.07	1
S11	72.14	2
S12	76.32	1
S13	79.85	3
S14	86.02	3
S15	88.75	3
Test	75	?

K=11 then test sample assigned to class2

Since mode is 5 for class2