Osteoporosis Risk Assessment with well-calibrated probabilistic outputs

Antonis Lambrou[1,2]
Supervisors: Harris Papadopoulos[1,2,3] and Alex Gammerman[2]
Advisor: Volodya Vovk[2]

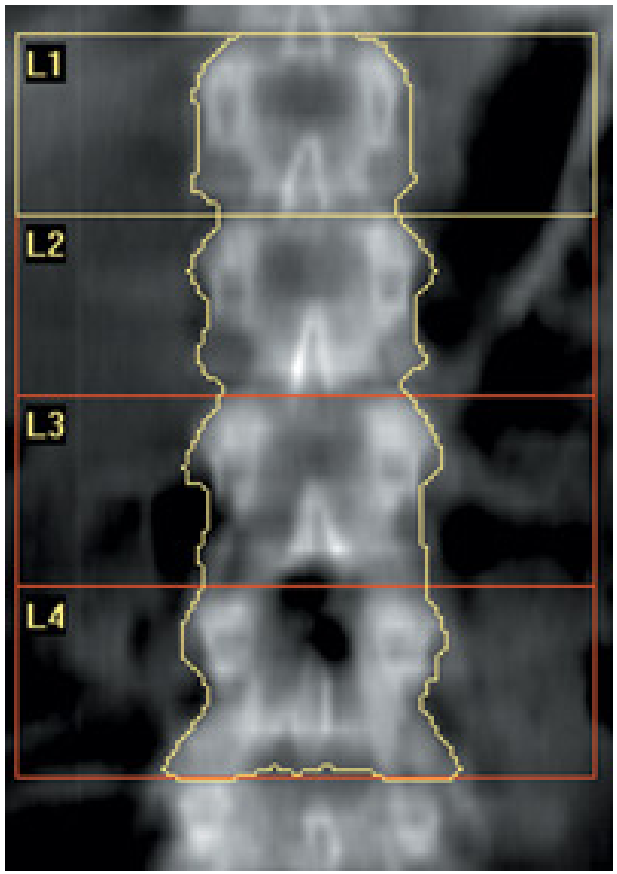
[1]Computer Learning Research Centre
Royal Holloway University of London

[2]Frederick Research Center, Nicosia, Cyprus

[3]Frederick University
Computer and Engineering Department Nicosia Cyprus

Problem definition

- Osteoporosis is a disease of bones that results in an increased risk of bone fracture.
- Dual-Energy X-ray Absorptiometry (DEXA) is used to diagnose Osteoporosis.



Problem definition

- We introduce the use of Venn Prediction in order to predict the risk of Osteoporosis before a DEXA scan, based on known risk factors.
- Unlike other probabilistic methods, Venn Predictors can provide well-calibrated probabilistic outputs under the assumption that the data used are identically and independently distributed (i.i.d.).

Venn Prediction

We calculate an empirical probability distribution for the true class of x_{l+1} based on the examples

$$\{(x_1, y_1), \ldots, (x_l, y_l), (x_{l+1}, Y_j)\}. \tag{1}$$

The Venn Prediction framework assigns each one of the possible classifications Y_j to x_{l+1} and divides all examples into a number of categories based on what is called a *Venn taxonomy*. For $n \in \mathbb{N}$, an n-taxonomy is a measurable function $K : Z^n \times Z \to \mathbf{K}$. The set $\mathbf{K}$ is usually finite; we will refer to its elements as categories.

After assigning the category $\kappa_i^{Y_j} = K((z_1, \ldots, z_l, (x_{l+1}, Y_j)), z_i)$ to each example in the extended set (1), the empirical probability of each classification Y_k in $\kappa_{l+1}^{Y_j}$ will be

$$p^{Y_j}(Y_k) = \frac{\left|\{i = 1, \ldots, l+1 | \kappa_i^{Y_j} = \kappa_{l+1}^{Y_j} \ \& \ y_i = Y_k\}\right|}{\left|\{i = 1, \ldots, l+1 | \kappa_i^{Y_j} = \kappa_{l+1}^{Y_j}\}\right|} \qquad (2)$$

So after assigning all possible classifications to x_{l+1} we get a set of probability distributions $P_{l+1} = \{p^{Y_j} : Y_j \in \{Y_1, \ldots, Y_c\}\}$ that compose the multi-probability prediction of the VP.

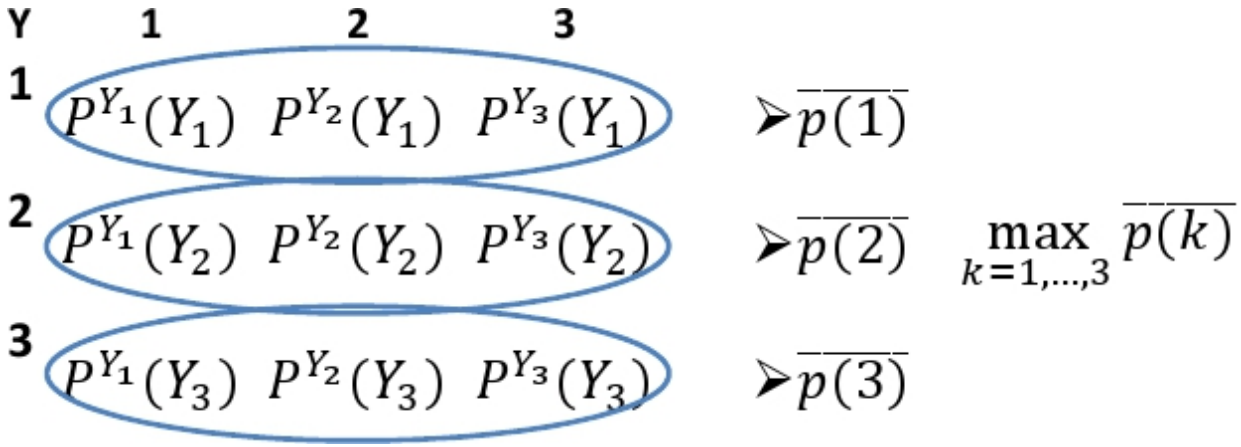
Venn Prediction

- The maximum and minimum probabilities obtained for each label Y_k amongst all distributions $\{p^{Y_j} : Y_j \in \{Y_1, \ldots, Y_c\}\}$, define the interval for the probability of the new example belonging to Y_k. We denote these probabilities as $U(Y_k)$ and $L(Y_k)$, respectively.
- The VP outputs the prediction $\hat{y}_{l+1} = Y_{k_{best}}$, where

$$k_{best} = \arg \max_{k=1,\ldots,c} \overline{p(k)} \tag{3}$$

and $\overline{p(k)}$ is the mean of the probabilities obtained for label Y_k amongst all probability distributions. The probability interval for this prediction is $[L(Y_{k_{best}}), U(Y_{k_{best}})]$.

Venn Prediction



Venn Predictor

- In this work, we use three underlying algorithms to specify three taxonomies: Sequential Minimal Optimisation (SMO), Random Forests (RF), and the J48 decision tree.
- Our taxonomies are based on the classification of the underlying algorithms.

Experiments

- We have collected 389 cases of female patients that have performed a DEXA scan.
- The data are constructed based on a questionnaire that is given to patients to complete.
- Each case is classified as "Normal" or "Risk of Osteoporosis" based on the patient's spine t-score that is given by the DEXA scan.
- According to the WHO, patients with a t-score above -1 are diagnosed as "Normal".

Experiments

#	Attribute name	Type	#	Attribute name	Type
1	Sex	Binary	35	Receive Thyroxine	Binary
2	Age	Numeric	36	Receive Estrogens	Binary
3	Weight	Numeric	37	Neurogenic Anorexia	Binary
4	Height	Numeric	38	Malabsorption syndrome	Binary
5	Start of Menstruation	Numeric	39	Chronic liver diseases	Binary
6	End of Menstruation	Numeric	40	Inflammatory bowel diseases	Binary
7	Pregnacies	Numeric	41	Transplantation	Binary
8	Smoking now	Binary	42	Chronic renal failure	Binary
9	Smoking in the past	Binary	43	Prolonged immobilization	Binary
10	No smoking	Binary	44	Cushing's syndrome	Binary
11	Years of past smoking	Numeric	45	Epilepsy	Binary
12	Years of current smoking	Numeric	46	Insulin Dependent	Binary
13	Cigarettes per day	Numeric	47	Ovariectomy before menopause	Binary
14	Alchohol intake per day	Numeric	48	Chronic gastrointestinal disorders	Binary
15	Caffeine intake per day	Numeric	49	Paget's Disease	Binary
16	History of fracture	Binary	50	Hyperthyroidism	Binary
17	Hip fracture	Binary	51	Parathyroid gland disease	Binary
18	Spine fracture	Binary	52	Receive Steroids	Binary
19	Wrist fracture	Binary	53	Receive Thyroxine	Binary
20	Low energy	Binary	54	Anticonvulsants (for seizures, epilepsy)	Binary
21	High energy	Binary	55	Diuretics	Binary
22	Sports	Binary	56	Heparin	Binary
23	History of osteoporosis	Binary	57	Chemotherapy	Binary
24	Osteoporosis in family	Binary	58	Treatment of osteoporosis	Binary
25	Loss of height	Binary	59	Alendronati	Binary
26	Kyphosis	Binary	60	Risedronati	Binary
27	End of menstrual bleeding	Binary	61	Zoledronati	Binary
28	Arthritis	Binary	62	Raloxifeni	Binary
29	Secondary Osteporosis	Binary	63	Strontio	Binary
30	Breast feeding	Binary	64	Parathormoni	Binary
31	Avoidance of milk	Binary	65	Denosoymapi	Binary
32	Avoidance of sex	Binary	66	Kalsitonini	Binary
33	Diarrhea	Binary	67	Calcium + Bitamin D	Binary
34	Receive Cortisone	Binary	68	Calcium	Binary

Table: Table of attributes in the Osteoporosis dataset.

Experiments

- We have performed offline 10-fold cross validation experiments on the data, with the J48 decision tree, Random Forests (RF), Sequential Minimal Optimisation (SMO), J48-VP, RF-VP, and SMO-VP algorithms.
- We also show the online results of the algorithms (training set grows as test examples are evaluated).

Experiments

Predictors	J48	RF	SMO	J48-VP	RF-VP	SMO-VP
Accuracy	70.18%	68.89%	67.10%	67.38%	65.17%	65.71%
Lower Probability	-	-	-	64.27%	57.93%	64.21%
Upper Probability	-	-	-	80.62%	78.09%	71.83%
Min probability $\geq$ 75%	-	-	-	54.73%	34.27%	7.61%
Min probability $\geq$ 70%	-	-	-	67.10%	60.51%	39.49%
Min probability $\geq$ 60%	-	-	-	75.53%	64.88%	84.52%
Accuracy at $\geq$ 75% min. prob.	-	-	-	73.61%	73.43%	74.05%
Accuracy at $\geq$ 70% min. prob.	-	-	-	72.71%	71.51%	69.01%
Accuracy at $\geq$ 60% min. prob.	-	-	-	72.06%	71.44%	67.93%

Table: Results of the six algorithms on the Osteoporosis dataset.

Experiments

We conduct experiments in the on-line mode: Initially all examples are test examples and they are added to the training set one by one after a prediction for each one is made. We graph the Cumulative Lower Accuracy Probability (CLAP), the Cumulative Upper Accuracy Probability (CUAP), and the Cumulative Accuracy (CA) curves:

$$CLAP(t) = \frac{1}{t}\sum_{i=1}^{t} U_i(Y_{k_{best}}), \tag{4}$$

$$CUAP(t) = \frac{1}{t}\sum_{i=1}^{t} L_i(Y_{k_{best}}), \tag{5}$$

$$CA(t) = \frac{1}{t}\sum_{i=1}^{t} Acc_i, \tag{6}$$

where t is the number of test examples that have been added to the training set, and $Acc_i = 1$ when the prediction for example x_i is correct and 0 otherwise.

Experiments

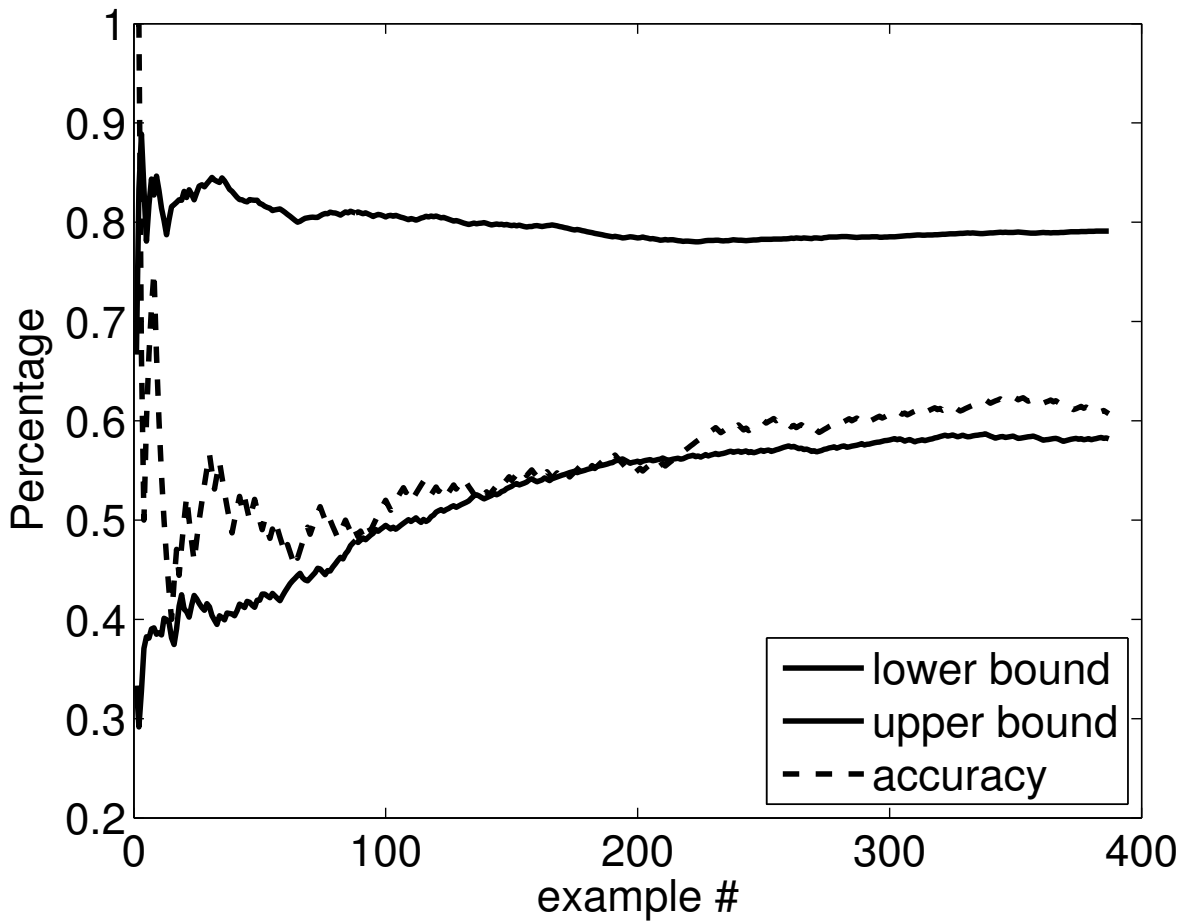


Figure: Online experiments with J48-VP on the Osteoporosis dataset.

Experiments

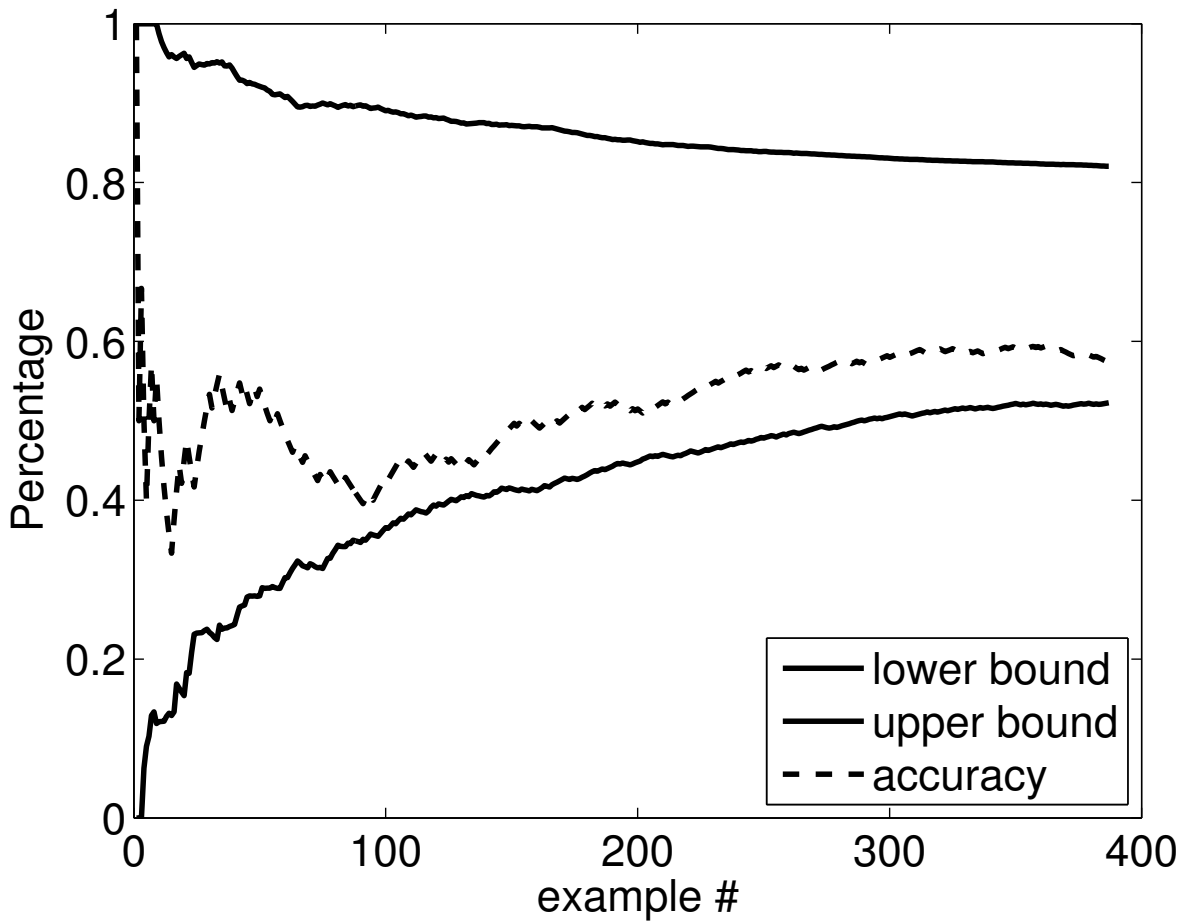


Figure: Online experiments with RF-VP on the Osteoporosis dataset.

Experiments

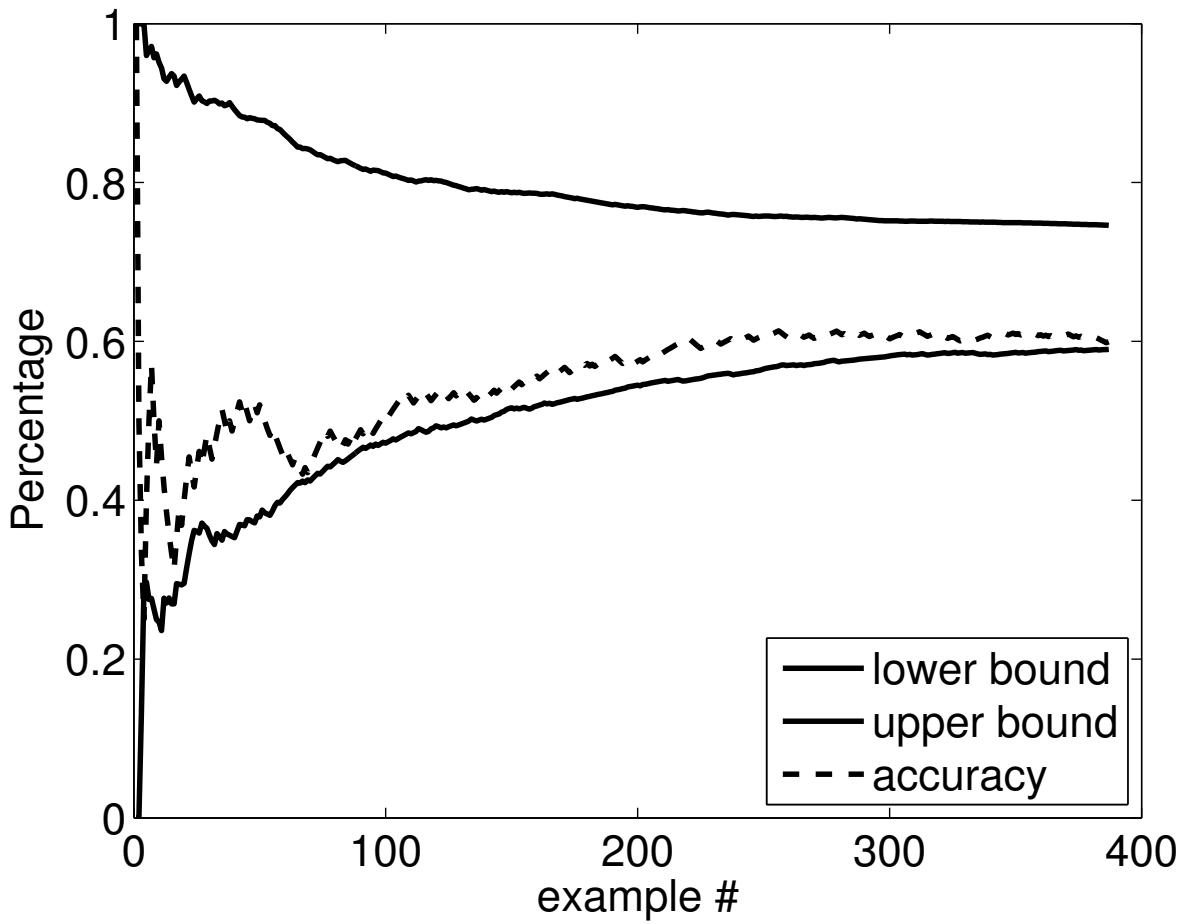


Figure: Online experiments with SMO-VP on the Osteoporosis dataset.

Conclusion and future work

- We have evaluated our method on real-world data that we have collected from various clinics in Cyprus.
- Our results, demonstrate that our method provides well-calibrated probabilistic outputs in the predictions that can be useful in practice.
- In the future, we aim to collect more data and perform supplementary analysis, in order to improve and evaluate further our VPs.
- Furthermore, we are in the process of building a tool for physicians that will enable them to use our VPs for assessing the risk of Osteoporosis.

Acknowledgement

- This work was co-funded by the European Regional Development Fund and the Cyprus Government through the Cyprus Research Promotion Foundation "DESMI 2009-2010", research contract TPE/ORIZO/0609(BIE)/24 ("Development of New Venn Prediction Methods for Osteoporosis Risk Assessment").

Questions

Thank you for your attention.

email: A.Lambrou@cs.rhul.ac.uk