MAY BAYESIAN CLASSIFIERS DIAGNOSE SCRAPIE?

Ingravalle F.*, Maurella C.*, D'Angelo A., Bona MC.*, Possidente R.*, Iulini B.*, Caramelli M.*, Ru G.***

**CEA - Centro Encefalopatie Animali - TSE National Reference Laboratory Istituto Zooprofilattico Sperimentale di Piemonte Liguria e Valle d'Aosta - Italy,*
***Faculty of Veterinary Medicine, University of Turin, Italy.*

Introduction: Scrapie is a neurological and fatal disease which affects sheep and goats and it is included in the Transmissible Spongiform Encephalopathies group. It is not easy to diagnose it when the animal is still alive as currently there are no *ante mortem* diagnostic tests able to detect it. Nevertheless it shows some quite peculiar clinical features. The mandatory report of scrapie suspects is a very important step in the struggle for its eradication.

Aim of this study is to provide a list of diagnostic signs (features) that have the greatest discriminatory power to categorise sheep into mutually exclusive scrapie and non scrapie groups. Usually a list is a necessary step to construct Bayesian classifiers to estimate the likelihood of being scrapie affected or not in presence or absence of certain features.

Materials and Methods: 282 sheep coming from scrapie outbreaks have been visited during a four-year period. On the basis of a standardise clinical examination, the symptoms were recorded and entered into an *ad hoc* database. The work is based on the assumption of accuracy of the features and their non total independence. However, even if independence is not completely verified, Bayesian classifiers have been proved to tolerate various degrees of dependencies [4]. The presence or absence of n=29 features (denoted by x_i , where i={1,…,n=29}) and the presence or absence of disease (***scrapie***=1 if the animal was affected by Scrapie, ***scrapie***=0 otherwise) were recorded. The accuracy of clinical signs was already evaluated in a previous study [2].

Hence, we classified features, single or grouped, following a Bayesian criterion:

- firstly: calculating the conditional probabilities for the individual features x_i:

$$P(x_i = 1 \mid scrapie = j), \text{ where } j=\{0, 1\}$$

- secondly: calculating the probability for the minimum classification error (***Pe***) for a general subset ***B*** of ***d*** features (***d***=1,…,***n***=29 is the subset size). This calculation was performed as reported in [3].

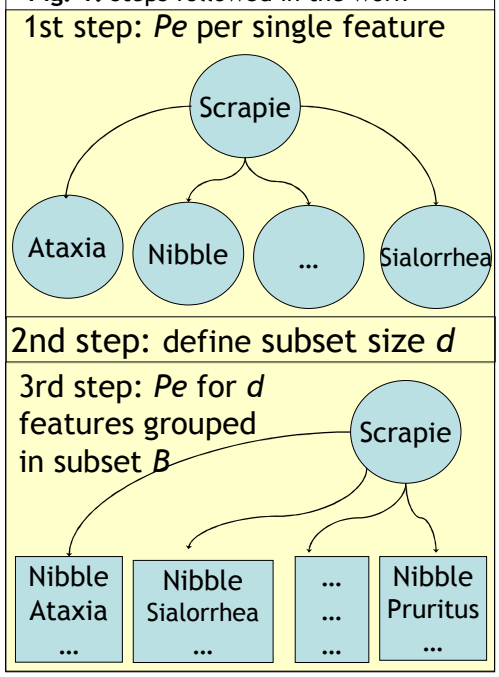
Then we followed these steps (fig. 1):

1) we estimated ***Pe*** for each of the ***n***=29 investigated features separately; ***Pe*** were sorted in ascending order (fig. 2).

2) we defined the size of the subset in ***d***=12. Therefore the subset size *d* is crucial. Usually the choice of *d* depends on two possible criteria:
   - one concerning the probability of error (i.e.: ***Pe*** lower than a fixed value);
   - one strictly concerning the features of the disease (i.e.: the most sensitive clinical signs or clinical signs easy to identify on field). This was our choice in the study.

3) a subset ***B*** was chosen in order to minimize ***Pe***, applying two methods: the single best method (SB in fig. 3) and the sequential forward selection method (SFS in fig.4)

Fig. 1: steps followed in the work



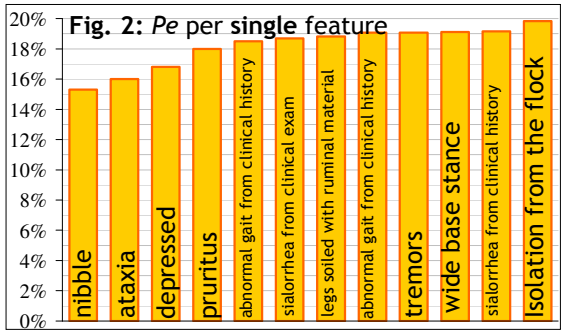
Results and discussion: the results show:

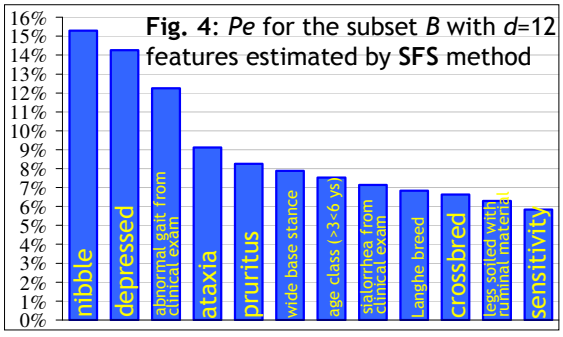
a) the bigger the information (i.e. larger *d*), the smaller the ***Pe*** (fig. 2 vs fig.3-4).

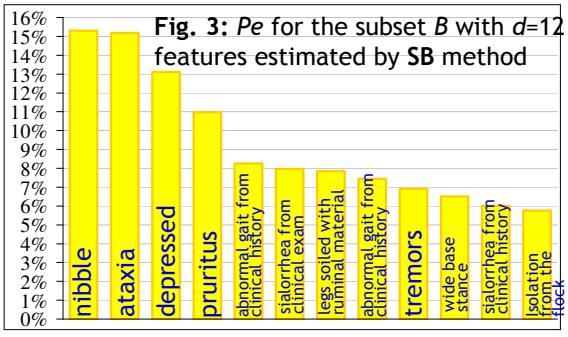
b) applying SFS method instead of SB method subsets with same size but different features are created. These methods don't assure the minimum ***Pe***, but they are acceptable compromises between efficiency and complexity of calculation.

c) the choice of ***d***=12 is related to the most sensitive clinical signs of Scrapie.

Moreover, data included in the study come from the on field practice and refer to individual clinical signs. On this basis, it could be useful to give out these results to the practitioners to enforce mandatory Scrapie reports.







References:

1. Geenen P.L., *et al.*: "Naïve bayesian classifiers for the clinical diagnosis of classical swine fever", proceedings SVEPM (2005);
2. D'Angelo A., *et al.*: "Assessment of clinical criteria to diagnose Scrapie in Italy",Vet J (2006, ahead of print);
3. Kuncheva L., *et al.*: "Selection of independent binary features using probabilities: an example from veterinary medicine",Journal of modern applied statistics methods, (2005), 4(1);
4. Domingos P., *et al.*: "On the optimality of the simple Bayesian classifier under zero-one loss",Machine Learning, (1997), 29: 103-130 .

SVEPM Annual Conference March 28-30, 2007

Dipoli, Helsinki/Espoo, Finland