

A Model for Rating and Certifying Competences in the EUCIP Framework

Andrea Violetti¹, Susanna Daddi², Stefano Hajek³

¹ AIP-ITCS

² Jobnet SpA

³ Perugia University, Insubria University

Introduction

Italian Computer Society - acronym AIP-ITCS - is a trade association which allows its members, through a “network of social relationships”, to communicate with and between different environments with the purpose of increasing the intellectual capital of computer professionals and the social capital of the association.

While waiting for better policies for innovation and ICT to be identified, we are aware that an “objective” and “merit based” management of knowledge, although difficult to achieve, should constantly be pursued, in opposition to the natural tendency of man to develop a “subjective” and “patron-and client” management of knowledge. We consider a merit based system the ideal system; in practice this means a constant tendency towards a merit based management.

The guideline for policies of social and economic development will therefore be a merit based and corporative rewarding welfare applicable to the “market of knowledge.” This is how the model for measuring and certifying competences originates. The aim is to achieve, in an objective and merit-based manner, a system for the assessment of competences, in this case those concerning the ICT sector, without leaving out, next to distinctive competences, relational and communicative skills.

The purpose of this model is to elaborate a standard method for the issuing of competence certificates in accordance with the EU directive 92/51 of 18 June 1992 and subsequent modifications and with the national regulations of reference, in order to show the market that a certain professional is competent in a specific or general profession and that that competence can be used all over the EU.

This model was first conceived in July 2005 with the purpose of finding a strategic and merit-based solution that fulfils the need for certification of competences in all intellectual jobs.

It has been necessary to reduce the complexity of the issue through the implementation of a model easily understandable and usable by the work, training, service, industry, trade and consumer markets and in all cases by any interested organization.

Basically, the idea is to calculate the average quality index of a professional organization and the position of the professional above or below this average, obliging those who wish to receive a competence certificate as general computer professional, to keep their index above the average through life-long training. Next to this we can identify specific profiles for which it is possible to obtain a further specialist competence certificate by verifying, through matching, the competence achieved in a specific sector.

1. Model for measuring and certifying competences

The model is a simplified and stylized representation of the reality which is being analysed, i.e. the measurement and certification of ICT competences; all elements considered irrelevant have been eliminated in order to focus on the essential through verbal, graphic and mathematical descriptions.

The model should be read in two ways:

as a description of what happens given the real values of the “exogenous elements” the results obtained for the “endogenous elements”	(measurement and certification of competence) (titles, experience) (which certification)
as an instrument for deciding what is to be done given the desired values of the “endogenous elements” values that the “exogenous elements” gain	(life-long training and self assessment) (in what I would like to evaluate myself/certify) (what I should study)

Meets

- The requests, standards and guidelines of the CEPIS
- Is compatible with the Eucip frameworks
- The law requirements and best practices of the sector
- The expectations of professionals, clients and enterprises
- Market expectations in terms of professional ethics and deontology
- Requests of both private and public training structures
- Practicality and functionality criteria
- Other professional associations

Supplies

- General criteria for the development of a data base of training paths in order to manage the relative credits
- General criteria for the development of a data base of professional competences in order to manage the relative credits

- Objective basis for issuing the competence certification in accordance with the EC directive 92/51
- A simple and objective classification of the competences of members
- Registration of the life-long training activities of the single members

The model results from the considerations of thousands of members during the last 15 years and does not expect to be conclusive or exhaustive but wants to make up for a serious deficiency in the ICT sector, i.e. the one of competence evaluation. Every association must be able to plan an independent and efficient implementation of the model in order to manage, measure and certify the competences of its members in the interest of professionals and of the market. Every training and professional activity must be adequately supported by computer instruments updated on –demand on training needs in order to overcome the complexity and heterogeneity that the society of knowledge interprets, and professionals must guarantee a minimum standard both with regards to knowledge (training) and to competence (experience), considering the paradigm: knowledge + experience = competence.

Considering that, as by article 8A of the EU treaty, the internal market means an area without internal barriers and that, pursuant to article 3, letter c) of the treaty, the elimination of barriers between member States for the free circulation of people and services is one of the objectives of the Community; that for the citizens of the member States this elimination implies the possibility of carrying out a profession, independently or dependently, in a member State different from the one in which they acquired their professional qualifications; considering moreover that according to the definitions specified in letter c) article 1, Title 1 of the EU directive 92/51, “competence certificate” means any title:

- ratifying a training path not part of a system constituting a diploma according to the EEC directive 89/48 or a diploma or certificate according to the present directive, or
- issued after an evaluation of the qualifications, aptitudes or knowledge of the requesting party considered essential for carrying out a profession from an authority appointed in accordance with the legislative or administrative regulations of a member State, without the request of the demonstration of a preliminary training;

The competence certificate issued with the aid of the above model has become fundamental to operate on the markets.

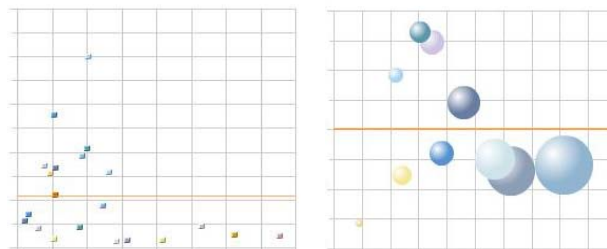
1.1 Certification of competences

The model for Measuring and Certifying Competences is preparatory to the issuing of the competence certificate in accordance with the EC provision 92/51 and subsequent modifications. The certificate is issued if the applying candidate’s final evaluation is above the average (considering a margin of tolerance for

acceptation) of all the final evaluations of the members of a professional association; this certification is general and certifies the professional ability of the candidate in the field of computers. After the issuing of the general certificate the professional can apply for a supplementary specialist certificate. The dynamism both of the data base of the candidates (on which the model is based), profiles, titles and certificates supplies a valid support both for the planning of life-long training and for the constant assessment of the “validity” of the granted certificate.

Compelling the candidate to remain always within the range complying with the minimum quality index automatically calculated.

As the evaluation is acceptable only if the average exceeds the minimum accepted level, the second candidate will necessarily have to be superior to the previous. This creates a mechanism of self-control on the skill status and professional capacities of the candidate and subsequent candidates. For this reason the model does not pay too much attention to the obligatory or obvious education that a professional ought to have, but focuses on the ability of the professional to keep himself updated and remain in the average of this organization which, in its turn, is forced to declare to the market its objectives in terms of average quality index of its members by publishing these indexes on an official web site. Depending on the organization’s and professionals’ requirements, the publication of the evaluation results for the certification of the candidate can be public, private or partially public; nevertheless they will be indicated on the competence certificate which can expire or be suspended if the candidate’s indexes of quality should go below the levels of non-acceptability indicated by the organization. This dynamic system forces the professional to follow constant training and to record his training; only through supplementary training hours and further titles or certificates can the depreciation of knowledge be compensated and constant improvement be generated with the introduction of new professionals with further qualifications and skills. All this applies to the minimum and general training of the professional; regarding specialist training, starting from 100 for the highest qualification in a specific field, the professional can be certified with 100 or below 100. The representation of the general profile is a dot situated on a Cartesian axis where the ordinates represent the RPC and the abscissas time (the age of the professional), while the graphic representation of the specialist profile is represented by a circle whose centre is situated where the general profile is located and the diameter representing the total coverage of the profile declared with a self-certification of skills.



1.2. Distinctive elements and definitions

<i>Competence certificate (ACK)</i>	<i>Issued by the association The competence certificate is a certificate issued after an evaluation of the personal qualifications, aptitudes or knowledge of the applicant considered fundamental for carrying out a profession by an authority appointed in accordance with the legislative and administrative regulations of a state. It rests upon the solid basis of life-long training, professional experience and deontology, providing an overall evaluation of the professional. The professional's competences will therefore be enhanced and evaluated differently. The market will finally be able to assess the true value of the professional to whom it has turned.</i>
<i>Relative Professional Credits (RPC)</i>	<i>Assigned to each item of the scientific committee in relation to computer sciences.</i>
<i>Relative Training Credits (RTC)</i>	<i>Number of hours of constant and compulsory training necessary each year to overcome deficiency of competences and relative professional credits (RPC)</i>
<i>Relative Deontological Debt (RDD)</i>	<i>Assigned by the industrial tribunal in relation to computer sciences (it must tend to 0)</i>
<i>Overall evaluation of the member (OEM)</i>	<i>Evaluation of the member from the scientific committee by means of an interview to assess, among other things, relational and behavioural skills.</i>
<i>Coefficient of devaluation or re-evaluation of the competence (CDR)</i>	<i>It is an indicator of devaluation or re-evaluation, therefore negative or positive, which assesses how much the value of the RPC/RTC given to a certain item decreases or increases in the course of the year.</i>
<i>Bonus malus index (BMI)</i>	<i>Positive and negative corrective index of the RPC/RTC at the disposal of the overall evaluation group of the member</i>
<i>Average quality index (AQI)</i>	<i>Average of all the overall evaluations of the professional (OEM), index above which the professional is associable and below which further training or certifications are required for the preservation or renovation of the certification</i>
<i>Percentage departure quality (PDQ)</i>	<i>Indicates how many percentage points of departure from AQI are accepted</i>
<i>Index Professional Quality (IPQ)</i>	<i>Index resulting from the application of the model; it represents the index of quality achieved by the professional</i>

1.3. Level of competences

The levels set by the model are two:

General Computer Professional

All professionals with an average RPC above the average index of the organization (AQI), represented on the graphic by dots.

Specialised Computer Professionals

All professionals with $RPC > AQI$ and 100% matching with one of the assessed profiles

(Example: Eucip Elective profiles – Business Analyst), represented on the graphic by circles.

Thanks to the very nature of the model, based on a data base and with a numeric measuring system of competences in a specific profession, if the simplification should prove to be insufficient, it is possible to elaborate further levels of competences and further competence certifications.

1.4. Data base

	% Incident of Section	Coefficient of devaluation or re-evaluation of the competence (CDR)	Relative Professional Credits (RPC)	Bonus Malus Index (BMI)	Relative Professional Credits (RPC) Calculate
Training Section					
Certification Section					
Skills section					
Publication Section					
Association Section					

Index Professional Quality (IPQ)

The model includes the management of Curriculum Vitae through the acquisition of information divided in sections, with a unit of measurement, weight, devaluation or re-evaluation coefficient for a positive or negative correction and a criteria for the assignment of RPC for each section.

1.5. Implementation level

The model has already been implemented in our association, has already lived its first cycle and has passed the four processes of innovation showing low levels of difficulty in the first phase, medium in the second and high in the tow final phases; now it is important to identify accelerators for all processes and to improve what has been achieved up to date.



2. Bayesian neural networks in human resources pattern recognition

Analysis of correspondences between job requirements and personal skills plays a central role in Human Resources Management; lacking skills can be filled by addressing training programs to people “potentially” interested in.

A way to discover professional potentiality is to explore skills patterns, making inference about the relevance of each skill “given” other skills.

Neural Networks are pattern recognition tools we used to investigate skills’ reciprocal support; instead of traditional maximum likelihood based learning we adopted Bayesian learning to avoid overfitting shortcoming: forecasting is integrated all over weight posterior distribution.

Being involved integration, analytically unfeasible, Hybrid Metropolis Algorithm was adopted; it numerically treats complex integrals making use of net gradient information to speed Metropolis random walk runs.

Real case simulations are commented with focus on approach’s robustness.

2.1 Job profiling

The case study we’re treating refers to iJobnet, a framework for human resource management that classifies professional skills, collects skills into a database and measures the fit of each skills aggregate (cv items) to ideal “job descriptions”; fit degree is a measure of distance between singular cv’s items and items in job descriptions.

iJobnet’s taxonomy for IT professionals follows EUCIP international standard; it is based on expert rules periodically updated by a scientific committee.

A fundamental iJobnet’s mission is to reduce skill supply - demand mismatching through development of targeted training programs to IT professionals; so iJobnet explores learning needs in that professional community to improve profiles’adequacy . This is accomplished by discovering recurrent skills patterns in the community and “completing” patterns in individuals: basic idea is to “fill in” lacking skills in profiles belonging to common knowledge areas (patterns), addressing training sessions to people “potentially” interested in.

Aim of this work is to design a pattern recognition tool that makes inference about the support of groups of skills to a singular skill so to evaluate the relevance of a skill given other skills.

2.2 Neural Networks and pattern recognition

2.2.1 Neural Network paradigm

Well established tools for pattern recognition are neural networks.

Neural networks are non-linear regression models inspired by natural metaphor of brain structure in which information is represented and stored through activation maps in a set of connected nodes.

Formally, define the node

$$y(\bar{x}_{i,l-1}, \bar{w}_{i,l-1}) = x_{i,l} = \left(1 + e^{-\sum_{j=0}^{\omega} w_{i,j,l-1} x_{j,l-1}} \right)^{-1}$$

Where $x_{i,l}$ is the i^{th} node at layer l , $\bar{x}_{i,l-1}$ is the vector of nodes $x_{j,l-1}$ at layer $l-1$, $\bar{w}_{i,l-1}$ is the vector of weights $w_{i,j,l-1}$ at layer $l-1$ “linking” $x_{i,l}$ with $x_{j,l-1}$ ¹

Being the node x defined recursively, the weighted sum of nodes along consecutive layers gives the overall network value for a set of inputs \bar{x}_0 .

Network’s learning is a “reinforcement” process that updates connection weights according to the minimum distance between observed examples and model’s representations.

Let Err be the net training error over a set of observation patterns: learning consists of training error reduction:

$$Err = \sum_{n=0}^N \left[\sum_{i=0}^{\omega} x_{i,l} - t_{i,n} \right]^2 = \sum_{n=0}^N \left[\sum_{i=0}^{\omega} y(x_{n,l-1}, w_{i,l-1}) - t_{i,n} \right]^2$$

being n ($0 \leq n \leq N$) the number of observations patterns and i ($0 \leq i \leq \omega$) the number of net output.

Traditionally training error reduction is obtained via least squares, hence reaching the maximum likelihood of observed data given parameters. [D.E. Rumelhart et Al - 1986].

This approach to inference is well known to be affected by overfitting: the training sample could be not representative of the full population’s behavior and too detailed models have the shortcoming to tradeoff a low in-sample approximation error with high variance – i.e. not acceptable out-of-sample (prediction) error.

2.2.2 Bayesian Neural Networks

Maximum likelihood criterion is reliable only when observed samples are the most likely; it is the most probable circumstance, but it is not certain. Conversely, a way to take into account parameters’ uncertainty is to consider the maximum

¹ The choice of sigmoidal function as “activation function” ensures neural network’s “universal approximation” property as stated in [K.Hornik et Al. 1989].

probability of parameters given data, parameters themselves following a distribution law.

Bayesian learning is based on this idea deriving posterior parameters distribution from the Bayes rule:

$$P(w|x, y) = \frac{P(x, y|w)P(w)}{\int P(x, y|w)P(w)}$$

In neural networks frame bayesian approach integrates predictions from all possible weights vectors over the posterior weight distribution rather than use a single “optimal” set of network weights.

With integration a weight vector that fits the data only slightly better than others will contribute only slightly more to prediction without exclusion of alternative models [R.M. Neal - 1992]:

$$\hat{y}_{n+1} = \int y(x_{n+1}, w)P(w|x_1, y_1, \dots, x_n, y_n)dw$$

Posterior distribution evaluation requires expressions for prior distribution and for likelihood function. Assume for the prior a zero-mean Gaussian function for the weights of the form:

$$p(w) = \frac{1}{Z_w(\alpha)} \exp \left(-\alpha \frac{1}{2} \|w\|^2 \right)$$

In which the normalization factor $Z_w(\alpha)$ is given by:

$$Z_w(\alpha) = \frac{\left(\frac{2\pi}{\alpha}\right)^W}{2}$$

Where W represents the total u of weight parameters. [C.M. Bishop - 1995].

Likelihood function represents here the distribution of target values for a given input vector; again we consider a Gaussian with mean given by the output $y(x, w)$ of the network and variance governed by a parameter β^{-1} so that:

$$p(t|x, w) = \left(\frac{\beta}{2\pi}\right)^{1/2} \exp\left(-\frac{\beta}{2} [y(x, w) - t]^2\right)$$

Looking at the entire dataset D we assume that the patterns are drawn independently from this distribution, and hence that joint probability is multiplicative:

$$p(D|w) = \prod_{n=1}^N p(t_n|x_n, w)$$

$$p(t|x, w) = \frac{1}{Z_D(\beta)} \exp\left(-\frac{\beta}{2} \sum_{n=1}^N [y(x_n, w) - t_n]^2\right)$$

Where the normalization factor $Z_D(\beta)$ is given by:

$$Z_D(\beta) = \left(\frac{2\pi}{\beta}\right)^{N/2}$$

2.3 Montecarlo numerical integration

Integral involved in Bayesian prediction is generally hard – if not impossible - to evaluate analytically; numerical methods, namely Montecarlo methods, are available for this purpose.

The fundamental idea is to approximate $\int y(x, w)p(w|x, y)$ with the

$$I = E[y(x, w)] \approx \frac{1}{N} \sum_{i=1}^N y(x, w_i)$$

corresponding expected value

where $\{w_i\}$ represents a sample of weight vectors generated from the distribution $p(w | x, y)$.

2.3.1 Metropolis Algorithm

This basic computation is for Bayesian neural networks infeasible due to the multi – dimensionality and multi – modality in the posterior probability density for w ; Metropolis algorithm is an effective way to arrange for the distribution of weight vectors to correspond to $p(w | D)$.

Metropolis Algorithm generates candidate steps from Markov chains:

$$w_{\text{new}} = w_{\text{old}} + \varepsilon$$

but reject a proportion of the steps which lead to a reduction in the value of $p(w | D)$.

This must be done with great care, however, in order to ensure that resulting sample of weight vectors represents the required distribution. In the Metropolis algorithm this is achieved by using the following criterion:

If $p(w_{new} | y, x) \geq p(w_{old} | y, x)$
Then accept

Else accept with probability
 $\frac{p(w_{new} | y, x)}{p(w_{old} | y, x)}$ ²

2.3.2 Hybrid Montecarlo algorithm

A critical point in Metropolis Algorithm and in any other MCMC samplers is the number of initial steps until the chain approaches stationarity; being that number related to the distance of starting value from the distribution's mode, the length of so named "burn-in period" can be reduced using back-propagation (gradient descent) to find starting values.

Hybrid Montecarlo method improves efficiency in Metropolis algorithm taking account of information concerning the gradient of $p(w | D)$ and using this to choose search directions which favour regions of high posterior probability. [D. MacKay - 1999]

It is inspired by energy evolution dynamics in physical systems as described by Hamiltonian equations.

Set $U(x) = -P(w|x, y)$; willing to draw Montecarlo Samples from $\pi(x) \propto e^{-U(x)}$ introduce a fictious "momentum vector" p with corresponding

"kinetic energy" $K(p) = \sum_{i=0}^d p_i^2 / m_i$ where m_i represents the "mass" of components i .

System's total energy is: $H(x, p) = U(x) + K(p)$; sampling (x, p) from $\pi(x, p) \propto e^{P(w|x, y)} = e^{-H(x, p)}$ target distribution $\pi(x, p)$ is obtained from marginal distribution $\int \pi(x, p)$.

² Convergence proof for Metropolis algorithm shows that sample generation process attains stationarity at reached distribution; see: [N.Metropolis et Al - 1956]

Omitting theoretical motivations – far beyond the scope of present work - Hybrid Metropolis algorithm proceeds schematically as follows:

- Generate a momentum vector p' from the Gaussian distribution $\pi(p) \propto e^{-K(p)}$
- Reach a new configuration in phase space (x', p') running the following leapfrog algorithm:

$$\begin{aligned} p' &= p' - \frac{1}{2} \delta \frac{\partial U}{\partial x} \\ x' &= x' + \delta p' \end{aligned}$$

For $m = 1$ to $l-1$

$$\begin{aligned} p' &= p' - \delta \frac{\partial U}{\partial x} \\ x' &= x' + \delta p' \end{aligned}$$

End For

$$\begin{aligned} p' &= p' - \frac{1}{2} \delta \frac{\partial U}{\partial x} \\ K(p) &= \sum_{i=0}^n p_i^2 \\ H(x, p) &= U(x) + K(p) \end{aligned}$$

- Let $(x_{n+1}, p_{n+1}) = (x', p')$ with probability:

$$P = \min[1, \exp\{-H(x', -p') + H(x_n, p_n)\}]$$
and let $(x_{n+1}, p_{n+1}) = (x_n, p_n)$ with probability $1-P$

3. Application to skills pattern recognition

3.1 Network setup

We tested Bayesian neural network model accuracy on a sample of 200 curricula splitting the sample in two: a learning sample consisting of 150 curricula used to update the weights, and remaining 50 curricula used as validation set to evaluate model's generalization error.

A curriculum is, for our purposes, a collection of professional skills related to a specific technical area; in this experiment we cover IT development area feeding the network with 14 skills in input – each assuming value 1 in presence of the respective skill, 0 elsewhere - and one skill as target.

The network is structured – after trying alternative configurations searching for better bias / variance tradeoff - with a single hidden layer of 28 nodes; Hybrid Metropolis runned along 50.000 iterations reaching a stationary distribution.

3.2 Results and conclusion

Figure 1 depicts the in-sample error distribution originated by all models (weights vectors) applied to skill absence/presence forecasting.

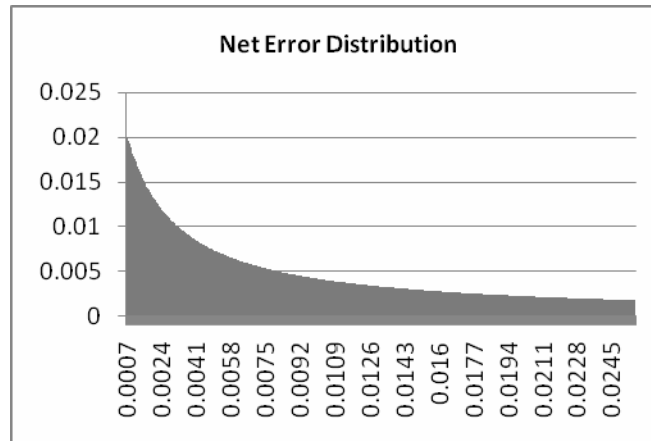


Fig. 1 – BNN error distribution

The network correctly predicts “in sample” the presence of a skill in the % of cases of effective presence and the absence in the % of cases of effective absence. Cross validation test (out of sample error) report a % right “presence” forecasting and a % right “absence” forecasting; this result confirms the good generalization property of Bayesian neural network especially when compared with results obtained with traditional approach (back propagation learning): as expected, while the in-sample error is lower in back propagation than in HMC, it becomes appreciably larger in back propagation than in HMC when measured out of sample (see table).

		False positive		False negative	
		HMC	BP	HMC	BP
In Sample	Out of Sample				

Bayesian learning reduce out of sample error taking into account the uncertainty of the model, hence integrating prediction with respect the posterior weights distribution.

This learning strategy is particularly interesting in situations – like this – where the learning sample is necessarily limited in size and not fully representative of the whole domain characteristics.

A further added value – only mentioned here - in Bayesian learning paradigm is the opportunity to incorporate in the model prior information from expert knowledge; opinions can be translated into prior distribution parameters value or directly or “forcing” the network to learn opinions: so obtained posterior distributions can be assumed as new priors in a new network.

Considered the central role of expert advisors in human resource analysis and selection this research direction will be seriously examined in the proceedings of our enquiry.

References

- Bishop C.M.: *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford, (1995).
- Hornik K., Stinchcombe M., White W.: “Multilayer feedforward networks are universal approximators”, *Neural Networks* Volume 2, Issue 5, p. 359 – 366 (1989)
- Metropolis N., Rosenbluth A.W., Rosenbluth M.N., Teller A.H., Teller E.: “Equations of State Calculations by Fast Computing Machines”. *Journal of Chemical Physics*, 21(6):1087-1092, (1953)
- MacKay D.: “Introduction to Monte Carlo methods.” *M. Jordan, editor, Learning in Graphical Models*, pages 175-204. The MIT Press, Cambridge, MA, USA, (1999).
- Neal R.M.: “Bayesian training of backpropagation networks by the hybrid Monte Carlo method”, Technical Report CRG-TR-92-1, Dept. of Computer Science, University of Toronto (1992)
- Rumelhart D.E., McClelland J.L., & the PDP Research Group: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1: Foundations*, MIT Press/Bradford Books Cambridge, MA (1986).