Expert Knowledge Makes Predictions More Accurate: Theoretical Explanation of an Empirical Observation

Julio Urenda[1,2], Marco Cardiel[2], Laura Hinojos[2],
Oliver Martinez[2], and Vladik Kreinovich[2]
[1]Department of Mathematical Sciences
[2]Department of Computer Science
University of Texas at El Paso, El Paso, TX 79968, USA
jcurenda@utep.edu, macardiel@miners.utep.edu, ljhinojos@miners.utep.edu
omartinez14@miners.utep.edu, vladik@utep.edu

Empirical observation that needs explaining. It is known that the use of expert knowledge makes predictions more accurate. A typical improvement – as cited in [1] on the example of meteorological temperature forecasts – is that the accuracy consistently improves by 10%. How can we explain this?

Towards an explanation. Use of expert knowledge means, in effect, that we combine an estimate produced by a computer model with an expert estimate. Let σ_m and σ_e denote the standard deviations, correspondingly, of the model and of the expert estimate.

In effect, the only information that we have about comparing the two accuracies is that expert estimates are usually less accurate than model results: $\sigma_m < \sigma_e$. So, if we fix σ_e, then the only information that we have about the value σ_m is that it is somewhere between 0 and σ_e.

We have no reason to assume that some values from the interval $[0, \sigma_e]$ are more probable than others. Thus, it makes sense to assume that all these values are equally probable, i.e., that we have a uniform distribution on this interval. For this uniform distribution, the average value of σ_m is equal to $0.5 \cdot \sigma_e$. Thus, we have $\sigma_e = 2 \cdot \sigma_m$.

In general, if we combine two estimates x_m and x_e with accuracies σ_m and σ_e, then the combined estimate x_c – obtained by minimizing the sum $\dfrac{(x_m - x_c)^2}{\sigma_m^2} + \dfrac{(x_e - x_c)^2}{\sigma_e^2}$ is $x_c = \dfrac{x_m \cdot \sigma_m^{-2} + x_e \cdot \sigma_e^{-2}}{\sigma_m^{-2} + \sigma_e^{-2}}$, with accuracy $\sigma_c^2 = \dfrac{1}{\sigma_m^{-2} + \sigma_e^{-2}}$. For $\sigma_e = 2\sigma_m$, we have $\sigma_e^{-2} = 0.25 \cdot \sigma_m^{-2}$, thus $\sigma_c^2 = \sigma_m^2 \cdot \dfrac{1}{1 + 0.25} = \sigma_m^2 \cdot \dfrac{1}{1.25} = 0.8 \cdot \sigma_m^2$, thus $\sigma_c \approx 0.9 \cdot \sigma_m$.

So we indeed get a 10% increase in the resulting prediction.

[1] N. Silver, *The Signal and the Noise: Why So Many Decisions Fail – but Some Don't*, Penguin Press, New York, 2012.