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

Julio Urenda\textsuperscript{1,2}, Marco Cardiel\textsuperscript{2}, Laura Hinojos\textsuperscript{2}, Oliver Martinez\textsuperscript{2}, and Vladik Kreinovich\textsuperscript{2}

\textsuperscript{1}Department of Mathematical Sciences
\textsuperscript{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
1. Empirical Observation That Needs Explaining

- It is known that the use of expert knowledge makes predictions more accurate.
- For example, computer-based meteorological forecasts are regularly corrected by experts.
- A typical improvement is that the accuracy consistently improves by 10%.
- How can we explain this?
2. Towards an Explanation

• Use of expert knowledge means, in effect, that we combine:
  – an estimate produced by a computer model and
  – an expert estimate.

• Let $\sigma_m$ and $\sigma_e$ 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 $\sigma_e$, then the only thing we know about $\sigma_m$ is that $\sigma_m$ is somewhere between 0 and $\sigma_e$. 
3. Towards an Explanation (cont-d)

• 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.

• So, we have a uniform distribution on this interval.

• For this uniform distribution, the average value of \(\sigma_m\) 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 \(\sigma_m\) and \(\sigma_e\),
  
  – then the combined estimate \(x_c\) is obtained by minimizing the sum \(\frac{(x_m - x_c)^2}{\sigma_m^2} + \frac{(x_e - x_c)^2}{\sigma_e^2}\).
4. Towards an Explanation (cont-d)

- The resulting estimate is \( x_c = \frac{x_m \cdot \sigma_m^{-2} + x_e \cdot \sigma_e^{-2}}{\sigma_m^{-2} + \sigma_e^{-2}} \), with accuracy \( \sigma_c^2 = \frac{1}{\sigma_m^{-2} + \sigma_e^{-2}} \).

- For \( \sigma_e = 2\sigma_m \), we have \( \sigma_e^{-2} = 0.25 \cdot \sigma_m^{-2} \), thus \( \sigma_c^2 = \sigma_m^2 \cdot \frac{1}{1 + 0.25} = \sigma_m^2 \cdot \frac{1}{1.25} = 0.8 \cdot \sigma_m^2 \), thus \( \sigma_c \approx 0.9 \cdot \sigma_m \).

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