Pattern Recognition Letters 64 (2015) 21–29

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

The disembodied predictor stance☆

Loizos Michael*

Open University of Cyprus, Cyprus

ARTICLE INFO

Article history:
Available online 30 May 2015

Keywords:
Introspective forecasting
Self-fulfilling prediction
Interconnectedness
Embodied cognition
Prediction recording
Unity of science

ABSTRACT

Pattern recognition is typically described as the discipline investigating how to recognize patterns and regularities in data, with the description leaving tacit that these patterns and regularities are somehow exploited, applied, acted upon, or simply announced once recognized. The aforementioned omission is more than a linguistic one, and is reflected on the emphasis that technical, theoretical, and empirical work on pattern recognition places on the predictors it develops, analyzes, and deploys. Most research on pattern recognition adopts, effectively, a stance amounting to treating the predictors as being disembodied, taken to mean that they operate without affecting the environment about which they make predictions. This essay argues for the dismissal of this stance, and demonstrates that the adoption of an embodied predictor stance is philosophically and technically not only possible, but also desirable.

1. Introduction

Much of modern science is intricately tied to the task of making predictions, broadly interpreted in this essay as the task of identifying the value of certain characteristics of an entity that are not observed themselves; at least not directly, or not at that particular point in time or space. One of the archetypical examples of prediction in the context of pattern recognition is visual recognition; e.g., predict what animal is depicted in an observed image, without being explicitly given that piece of information.

A reasonable assumption for many pattern recognition tasks, and certainly for the visual recognition task above, is that whatever the appropriate prediction is for a given entity, this value is determined independently of the agent itself. In particular, it is assumed that the agent's act of announcing a prediction does not affect what the appropriate prediction is. Such an assumption is suited more generally when making predictions in static or predetermined environments, where it is, even in principle, impossible to affect the appropriate prediction once the entity about which the prediction will be made has been observed.

There exist, nonetheless, pattern recognition tasks such that: *(C1)* an agent seeks to predict characteristics of a *future* state of the environment that has not yet materialized at the time of prediction, and which, in principle, can still be causally affected by the prediction; *(C2)* the agent has a *choice* on whether to announce its prediction and what prediction to announce, and it is thus meaningful to consider how to best go about doing so. Such prediction tasks include predicting the stock market, the outcome of an election, the final exam grade of a student, etc.

Despite the *theoretical possibility* in these scenarios that the act of announcing a prediction may critically influence the outcome one wants to predict, it would seem that most typical pattern recognition frameworks operate as if a predictor exists *outside* the environment. We suggest, in fact, that this presupposition is sufficiently widespread that one could view it as a *stance* in current pattern recognition research: predictors are treated as being *disembodied*, in that their output (i.e., the announcement of their predictions) affects only their mind, but not their body and, more generally, the environment in which they exist.

Fawcett [8], for instance, acknowledges the existence of such a stance in the context of spam filtering, suggesting that "data miners may have to consider the effects of mining on their task environment". He continues to note that considering such effects is still "foreign to most data mining researchers: the data are mined and the results are deployed, but the data environment is not considered to be an active entity that will react in turn".

This essay examines pattern recognition contexts where the announcement of a prediction by an agent can influence the outcome that the agent seeks to predict. We start by discussing in Section 2 whether dismissing the disembodied predictor stance altogether is warranted, and make explicit in Sections 3–5 the working hypotheses that would support such a change in stance. We continue by recounting in Section 6 the introspective forecasting framework [25] — a particular way to extend supervised classification models so that they acknowledge the effects of the predictions being made, on their

☆ This paper has been recommended for acceptance by Marcello Pelillo.
* Tel.: +357 22411963.
E-mail address: loizos@ouc.ac.cy

http://dx.doi.org/10.1016/j.patrec.2015.05.009
0167-8655/

realizability. Certain applications of introspective forecasting are then elaborated upon in Sections 7–9, before we conclude in Section 10.

2. Is a change needed?

In a time series forecasting setting [2], an agent observes a sequence of states, and seeks to predict certain characteristics of a certain subsequent state. In a first scenario, the agent might observe frames in a movie, and is expected to predict which actors appear in the next frame. In a second scenario, the agent might observe stock market transactions as they unfold in real time, and is expected to predict the price of a certain stock in the next transaction. Whatever the case might be in the next movie frame, the agent's announcement of its prediction will clearly not affect the outcome. It is evident that the same is not true in the stock market scenario, where both Conditions *(C1)* and *(C2)* are satisfied: the very act of the agent announcing a prediction might critically affect the characteristics of the next transaction by having investors react to the announcement.

What would it mean, then, for a prediction in the second scenario to be accurate? We shall adopt the most obvious answer that the prediction is still expected to match the stock price in the next transaction, by taking into account whatever effects and ramifications the announcement of the prediction may have. In the second scenario, therefore, one is forced to acknowledge the embodied nature of the predictor, in that it is an integral part of the environment for which it seeks to make predictions, and the announcement of its predictions might affect the environment and ultimately the realizability of the predictions themselves.

The embodiment of predictors in the described sense is not in line with the mainstream stance in pattern recognition. Do scenarios like the second one above offer a sufficient reason for the adoption of a different stance on the matter of the embodiment of predictors? Should we dismiss the view that predictors are disembodied, and explicitly and universally adapt our learning models to acknowledge the effects of their own predictions?

We continue to identify and present certain counterarguments to the adoption of a different stance. First, announcing or acting upon predictions in numerous cases does not appear to affect their realizability, *even* in scenarios where both Conditions *(C1)* and *(C2)* are satisfied. Every day, weather forecasters announce their prediction for tomorrow's weather, still there is no reason to believe that these announcements change the weather in any measurable way. Every day, people (attempt to) predict the lottery numbers and bet based on their predictions, still (assuming the process is not rigged) there is no reason to believe that their betting actions bias the winning lottery numbers. Second, even when announcing a prediction measurably affects the prediction's realization, it would seem that our existing pattern recognition techniques produce predictors able to anticipate and cope with the potential falsification of their predictions, as evidenced by the success of pattern recognition in numerous settings.

We could respond to these counterarguments by claiming that it is because of choices we make that we do not get to see predictions being falsified. But that response would not invalidate the essence of the counterarguments: the domains for which we wish to make predictions, the pattern recognition techniques we use to produce predictors, and the ways we choose to announce and act upon our predictions, happen to be such that predictions are not self-defeating. Thus, if by adopting the disembodied predictor stance — as we currently do — we can still get correct predictions, then there is no need to change that stance.

Our response would be that even if the disembodied predictor stance works well in current practice, we may gain a deeper understanding on *why* it should work by adopting the conceptually more accurate embodied stance. Further, adopting the embodied stance may allow the use of pattern recognition techniques in other domains as well, which have either not been explored before, or have been explored and abandoned due to the inapplicability of our existing disembodied predictor machinery. This last scenario suggests, in particular, the possibility of a hindsight bias in the counterarguments above: it might be that our current choices lead us to obtain useful predictors because we have stopped making those choices that had led us to obtain bad predictors. As a case in point, the technical analysis toward the end of Section 6 shows that the embodied stance can lead only to richer/more learnability results, since one can learn at least whenever learning is possible under the disembodied stance.

As a third counterargument, one may oppose the use of models that seek to ensure prediction validation, because in certain scenarios one explicitly seeks the opposite. A doctor predicting the illness of a patient acts upon her prediction to administer a treatment that explicitly aims to avoid the illness. Analogously, a teacher predicting the failure of a student to pass an exam based on past performance acts upon his prediction to assign extra course work with the explicit goal to avoid the failure.

This counterargument is easily rebutted by rewording it in a manner that fits the embodied predictor stance. A doctor may simply predict the wellness of a patient, with the action associated with this prediction being the administration of a treatment that would ensure the prediction's realization. Analogously, a teacher may simply predict the success of a student, with the action associated with this prediction being the assignment of extra course work that would ensure the prediction's realization.

One could come back with a fourth counterargument by opposing the use of any model that seeks to bring about the validation of its predictions, because of the self-fulfilling nature of the predictions to which the model will tend to give rise. However, self-fulfillment in itself is not problematic; it might become so when (and only when) it takes precedence over other metrics of success that are affected by an agent's predictions. A doctor's/teacher's concern (metric of success) is not to validate their predictions, but to achieve a beneficiary outcome for the patient/student. Making correct predictions is not an end in itself, but only a means, and prediction fulfillment should not take precedence over the realization of whatever the sought outcome is.

Admittedly, the introspective forecasting model that we shall later propose places primary emphasis on the fulfillment of predictions, and given an introspective forecaster's ability to affect the evolution of the world through prediction announcement, it will certainly announce those predictions that would maximize their fulfillment at the expense of any other concerns.

This brings up a central point in our essay, directly related to Condition *(C2)*: the distinction between the *agent* that chooses how to act — and which, if any, prediction to announce — and any pattern recognition *predictors* (possibly multiple ones) that the agent may decide to consult when making such a choice.

A pattern recognition predictor does not *choose* whether and what to predict; it predicts only as prescribed by its semantics and any given training material. If the pattern recognition predictor is, for instance, one that adopts the disembodied predictor stance, then its predictions will be accurate assuming that the agent will not announce them, or assuming that Condition *(C1)* does not hold. If the pattern recognition predictor is like the introspective forecasting one, then its predictions will be accurate if the agent announces the predictions (in the manner prescribed by the predictor). If both predictors used by the agent produce highly accurate predictions (under their respective working assumptions), the agent may rationally choose to announce the latter prediction if and only if the (self-fulfilling) prediction that was made is indeed more preferable to the former prediction.

By way of example, a teacher uses the introspective forecasting model and predicts that a certain student will score 63/100 in her final exam, if this introspective prediction is announced. The teacher uses also a typical learning model and predicts that the same student will score 76/100 in her final exam, if nothing is announced. In this