

7

THE FUTURE OF PREDICTION: FROM STATISTICAL UNCERTAINTY TO ALGORITHMIC FORECASTS

THE UNCERTAINTY OF THE OPEN FUTURE

The spread of learning algorithms is changing the meaning and forms of prediction, affecting the image of the future and the way to deal with it in the present. Whereas in the current view, the future is seen as open and unknowable because it does not yet exist and depends on present actions and expectations, today's predictive algorithms claim to foresee the future.¹ This claim is both exciting and frightening. It may lead to optimization of the use of resources and to targeted and effective prevention and planning, yet also may bind the future with preemptive policies based on existing patterns.² In any case, it breaks with the current idea of the future and management of uncertainty. My point in this chapter is that algorithmic prediction is very different from the idea of prediction that has established itself in modern society since the eighteenth century, oriented and guided by the calculus of probability; that is, it differs from the mathematical treatment of chance that began with the work of Blaise Pascal and Pierre de Fermat in the second half of the seventeenth century.³ Whereas probability calculus offers a rational way to deal with uncertainty,⁴ algorithms claim to provide an individual score for individual persons or singular events.

As studies on the emergence of statistics in the late seventeenth century show, forms of prediction change over time and have important consequences for society. When, as is currently happening, the forecasting agent is an algorithm and not a human being, processes and criteria are different, and results and problems change as well. Algorithmic prediction produces outcomes that would be impossible for a human being to generate, even if equipped with the tools of statistics; yet it also raises different problems that our society has to manage. This chapter aims to investigate these recent developments from a broader social perspective.

We'll see that while machine-learning systems are statistical engines, these systems and statistics are increasingly diverging. In fact, some algorithms, though products of the most advanced scientific practices, bear a surprising resemblance to some of the structures of the magical and divinatory mentality of ancient societies, which today are seen as directly opposed to science. Divination assumed that the future could be known in advance, even if human beings normally could not see it. For centuries instead, scientists of modern society have used statistical tools to manage the future's uncertainty. While machine learning inherits the tools of statistics, it tries, like divination, to foresee future events.⁵

DIVINATORY ASPECTS OF ALGORITHMIC PREDICTION

The task of algorithms is to predict the future. Amit Singhal, the former head of Google Search, explicitly stated this in 2013:⁶ that from now on, the primary function of search engines will be anticipating—predicting which information we will need rather than answering queries we have made. The objective of AI, claimed Kitchin, “is more to predict than to understand the world. Prediction trumps explanation.”⁷ Many projects that previously used digital tools for the purpose of managing information to explain phenomena now have turned to prediction.⁸ The goal of precision medicine, for example, is often to guide prognosis and effective treatment, even when the cause of the disease is still unknown. The move of algorithms from explanation to prediction, however, deeply modifies the meaning and the premises of prediction, together with the use of statistics.

Statistical methods can be used for causal explanation, as is currently the case in many areas of research, particularly in the social sciences.⁹ A theory suggests hypotheses that are tested with probabilistic tools. Statistics, however, can also be used for empirical prediction. In the first case the aim is finding the “true” model, while in the second case, the goal is finding the best predictive model, with the two goals failing to always overlap. Shmueli shows that in the practice of statistical modeling, the difference between “explaining” and “predicting” is often hidden by a common misconception: if one can explain, it is assumed one can predict.¹⁰ Predictive capability is subordinated to the ability to explain. Instead, the two are different and should be evaluated separately. In the use of models, the indiscrimination between explanation and prediction can lead to serious consequences. For example in the financial crisis of 2007–2008, economists and governmental agencies relied upon the capital asset pricing model (CAPM), which had been evaluated in terms of its explanatory power. But that capacity was not matched by its predictive power, which turned out to be far lower. This worsened the crisis in palpable ways.

Today, however, the availability of very high computing capacity and huge amounts of data generates new possibilities for using statistical tools primarily for predictive purposes. This does not mean, as claimed by some controversial positions in the debate about big data, that explanation has become superfluous and the search for causality obsolete.¹¹ Instead, it highlights the possibility and the need to distinguish the two goals and to analyze the scientific specificity of prediction, with its forms, its procedures, and its problems—that are different from those of causal explanation.

The modern scientific approach was developed in a time in which science aimed at explaining general results. Even if one can never apply a generalization from a specific finding to other, different cases (the philosopher David Hume’s classic problem of induction), probability calculus provides a stringent method and a rational basis for extrapolating from an inevitably circumscribed set of observations to a generalization about all cases.¹² Modern scientific procedures are based on a limited number of carefully selected data, the *experimental* data gathered during sampling that is processed to test the hypothesis formulated by a theory. Collecting

all data is not possible, and in the statistical approach, it is not even necessary because one only needs an appropriate sample large enough to be representative. The data, in a sense, are in the service of the theory, that is, serve to validate the hypothesis that explains the phenomena.

Digital procedures work differently in that they rely on enormous amounts of data and on sufficient computing capacity to manage them. Algorithms use *all* data that can be accessed,¹³ without “cleaning them up” to correct inaccurate or biased records and without selecting data points, which usually thereby include a myriad of secondary data collected for other purposes. Algorithms that recommend medical procedures, for example, not only use the patients’ medical records, but also data from their credit histories, from their relationships with acquaintances, or from their buying habits. The data in this case come before the theory, in the sense that a hypothesis, if formulated, is guided by them. One does not know what one is looking for, but sees what emerges from the data, which are largely unstructured. In the elaboration of the data, one does not look for causal relationships that confirm the hypothesis (because there is no hypothesis); instead, the search is for associations and correlations, for patterns whose detection discloses underlying structures and should make it possible to formulate effective predictions.¹⁴ On the basis of the patterns, one should be able to predict future developments, even if one cannot necessarily explain them. Predictive modeling differs from explanatory modeling.

Models that do not explain often cannot be explained, and the consequence is the much-debated nontransparency of algorithms.¹⁵ While the hypothesis guiding the explicatory approach must be understandable, in predictive modeling, transparency is of secondary importance: one should focus on “predictive accuracy first, then try to understand.”¹⁶ Algorithmic methods such as neural networks or random forests are often not interpretable, yet they make it possible to work with heterogeneous data and formulate effective forecasts. One can predict without understanding.¹⁷

Digital procedures are extremely innovative and their results are often astonishing. Yet a surprising aspect emerges: some features of the predictive use of machine-learning algorithms resemble an ancient, prescientific logic.¹⁸ The terms used in algorithmic prediction (“correlations,”

“patterns”), the idea of predictions independent of causal relationships, the reference to structures inaccessible to human reasoning, all have ancient and complex traditions in divinatory societies, as in the Middle East and Greece, and as developed in very elaborate ways in Chinese culture.¹⁹ Like algorithms, divinatory procedures were guided by precise techniques that rigidly provided a number of steps to be taken.²⁰ In both cases there are programs that, unlike scientific practices, do not attempt to explain or understand phenomena, but just try to deal with them.²¹

Like the procedures of machine-learning algorithms, the structures at the basis of divination in ancient times were obscure to the human mind.²² Divinatory societies relied on the assumption that the world was governed by a cosmic logic and by a basic order that human beings, with their limited capacities, were not able to grasp,²³ just as today we cannot fully understand the procedures of algorithms. Divinatory rationality was not of a scientific but of a ritualistic kind,²⁴ with the aim not of providing explanations but of managing a “total knowledge” that remained inaccessible.²⁵ As with algorithms, the goal was not to understand the phenomena but to get directions for action and decision.

The whole universe was taken as infinitely significant, articulated in an inexhaustible network of correspondences.²⁶ Just as the four seasons corresponded to the four compass points, and the history of a country to its topography, so the life of an individual corresponded to his or her body and his or her fate was inscribed in the order of things. The underlying *correlations* could be captured by identifying configurations and *patterns* in different phenomena: the walnut maple has the same shape as the human brain, the sky is the mirror image of the earth below it, the malformations in newborn humans resemble ominous terrestrial events. These phenomena were all “saying the same thing”;²⁷ therefore, by analyzing patterns in accessible phenomena, divinatory observers believed they could gain indications about correlated, inaccessible ones. From patterns in the liver of sacrificial animals or the flight of birds, with divinatory techniques, one could draw conclusions about the divine plans for the future and directions on the decisions to be taken, without understanding the reason or claiming to explain it.²⁸

For many nowadays, the idea that one can make decisions on the basis of the configurations of the liver of a lamb or of the starry sky seems

absurd; but the comparison with divinatory practices can enlighten the way in which current algorithmic predictive practices rely upon intricate, barely visible webs of connections.

MANAGING FUTURE UNCERTAINTY

Despite their striking structural similarities, a basic difference separates algorithmic and divinatory procedures: the underlying concept of time. When can a prediction be trusted and what does its credibility rely on? In the ancient divinatory worldview, the idea of anticipating the future was plausible because of the assumption that it was possible to see its structure in advance. The challenge was how. In various forms, in ancient times the basic distinction was between divine temporality and the temporality of human beings. In Mesopotamia the gods used signs to indicate future events to humans.²⁹ In Greece the gods were placed in eternity (*aeternitas*), while human beings were bound to time (*tempus*).³⁰ Seen from the divine perspective, the unknowable future appeared no less structured than the past, but human beings could not access it.³¹

In this ancient view, divination was rational, existing as a complex of procedures and techniques that made it possible to “give shape to an amorphous future.”³² To rely on oracles was not superstition and fantasy because of the assumption that the future had a structure already in the present, even if human beings could not know it. The indications one got from omens were uncertain, not because the relationship between the future and the present was uncertain, but because humans could not be sure to properly understand a higher perspective that as a whole remained inaccessible. Divinatory responses were enigmatic and required interpretation. If the verdict turned out to be incorrect, the interpretation was wrong, not the prediction.³³

This way of seeing time has its consistency and plausibility, but it is not that of the modern world nor of contemporary societies. Our concept of time presents the future as an open field, which today cannot be known either by humans or by any hypothetical superior entity, because it does not yet exist.³⁴ The future is not a given, but a horizon of the present that moves away as we approach it and can never be reached. What we can know about the future is not the future, but only the present

image of the future: our expectations and the information on which they are based. On the basis of these data, which exist and are observable, we can investigate and gather more detailed and reliable information. The prediction takes the form of *planning*: preparing the present to face in a controlled way a future that is always obscure. Because we cannot know in advance what will happen tomorrow, we calculate and manage our present uncertainty. Since the early modern age, the tool for dealing with the uncertainty of the future, instead of divination, has primarily been the *calculus of probabilities*.³⁵ The calculus does not promise to reveal what will happen tomorrow, but to calculate the probability of a particular future based on how much requisite knowledge we have now in the present (e.g., 40 percent or 27 percent). This enables us to decide something rationally even in face of uncertainty (i.e., even if things can disappoint our expectation).

The approach of statistics was developed in contrast to the divinatory tradition, and its empirical experimental approach became the basis of the scientific and technological attitude of modernity. Instead of interpreting signs, one gathers data; instead of discovering correlations, one notes empirical regularities. This approach was enormously successful, leading to the impressive development of scientific research. Now, however, this very research is producing the advanced techniques of machine learning and algorithmic prediction. These techniques, using statistical tools derived from probability calculus, can be used for prediction, thereby contradicting the assumption of the open, unpredictable future.³⁶ In ancient times the structure of the future appeared unknowable to humans but not to the gods; today the future appears to be unknowable to humans, yet should be accessible to algorithms.³⁷ How are algorithmic prediction and probabilistic tradition connected and distinguished?

AVERAGES VERSUS INDIVIDUAL PREDICTION

The key to the smartness of algorithms and all they can do, including make predictions, is the techniques that make it possible for machine-learning systems to autonomously develop the ability to process data and produce their own information. To do this, algorithms need examples of tasks to fulfill, and the web offers a lot of them. If a software program is

able to learn, those examples can be used to train algorithms in a more and more accurate and differentiated way. The diversity of contexts on the web becomes the resource for learning and increasing the performance of algorithms.

How do machines learn from examples? To develop this ability, the programmers in machine learning use the tools of statistics.³⁸ In fact, statistics and probability calculus addressed for centuries the problem of learning from data and produced a number of computational tools to extract information: regression, classification, correlation, and so on. Now machine learning inherits and adopts them, yet uses data in a different way. The goal of statistics is to manage present uncertainty. It addresses the knowledge (or lack of knowledge) of the present, maintaining and confirming the insuperable barrier between the present moment and the open future. Machine learning, instead, addresses the future and has the goal of predicting it. The difference between the two approaches produces a curious relationship of closeness and opposition between machine learning and the tradition of statistics, two formally almost identical cultures that are progressively diverging.³⁹ Even if they use the same tools, the attitude of machine-learning programmers is very different from that of statisticians, as their problems are different from the ones raised by the “avalanche of numbers” in the nineteenth century.⁴⁰

Statistics wants to contribute to knowing the world by activating a procedure that matches the classical Galilean method: inserting past data into the model and then using it to predict future data, thus verifying the accuracy of the model and eventually correcting it. The goal is explanation: when you do statistics, you want to infer the process by which the data were generated. For machine learning, on the contrary, the purpose is not to explain the phenomena elaborating a model. In many cases, you do not even know if there can be an intelligible model, and the machine can operate without one. The goal of algorithmic processing is not truth but predictive accuracy.⁴¹ In machine learning you start from the assumption that you are dealing with “complex, mysterious and, at least, partly unknowable” models.⁴² You do not want to understand them but to know how the future will look like with regard to some variables. Machine learning faces the future and tries to predict it as accurately as

possible, independently of our knowledge of the world. As we can read in a web debate, “statistics emphasizes inference, whereas machine learning emphasizes prediction.”⁴³

As a consequence of their different attitudes, statistics and machine learning produce fundamentally different forms of prediction. Statistics uses samples based on a limited amount of specifically prepared and selected experimental data in order to deal with the statistical universe. Statistics produces findings about the *average of the elements or subjects involved*—that is, results that correspond to nothing specific and to no one in particular (nobody has 1.4 children); however, these results increase our general knowledge. Algorithmic procedures, instead, use all available observational data and work with very large data sets, but produce no general results. They indicate what can be expected *for a specific subject at a given time* on the basis of correlations found in the data.

This feature of algorithmic procedures is similar to ancient divination, which also did not respond to an abstract interest in explanation but to a specific individual’s very practical questions: How should I (a particular individual) behave today to be in the most favorable condition tomorrow?⁴⁴ Where should the new city be founded? What is the best time to start a battle—or to sow wheat? Will my marriage be successful? The divinatory response produced punctual and individual predictions.⁴⁵ Likewise an algorithmic forecast is specific to the case before it. “Whereas forecasting estimates the total number of ice cream cones to be purchased next month in Nebraska, PA [predictive analysis] tells you which *individual* Nebraskans are most likely to be seen with a cone in hand.”⁴⁶

This is the main difference between the tradition of statistics and new developments in machine learning. Digital techniques abandon the statistical idea of averaging, in which all elements of a population represent more or less imperfect replicas of the average value.⁴⁷ The approach of big data claims to be more realistic because it rejects this abstraction and claims to process individual elements of the population with all their idiosyncrasies and incommensurability. The new frontier of customization will lie in the movement from the search for universals to the understanding of variability. According to the perspective of predictive analytics, “now in medical science we don’t want to know . . . just how

cancer works; we want to know how your cancer is different from my cancer. . . . Individualization trumps universals.”⁴⁸ Society is calculated without categorizing individuals, but by considering the specificity of everyone. Calculations start from people’s activities and do not try to infer features applicable to larger phenomena.⁴⁹

Paradoxically, the focus on individual specificity is achieved through neglect of the individual perspective, and actually of any perspective.⁵⁰ Algorithms should be able to predict the singularity of subjects because they do not depend on what people think and want, nor on what they say they want. Algorithms base their calculations on what people actually do, often without saying so or even without knowing it.⁵¹ What the algorithm treats as the perspective of the single individual is derived from digital “footprints” of people’s activities: zip codes, credit reports, driving records, language patterns, friends and relationships, and many other elements that are compared with similar data of other individuals.⁵²

But even if algorithms do not depend on a specific perspective, their personalized indications cannot be extended to other cases. They only apply to the available data set (with its implicit biases), to the targeted individual, and to the particular moment. That the results are local, specific and provisional, however, should be their strength. In the words of Andy Clark: “Context, it seems, is everything.”⁵³ Learning algorithms are extremely effective and can achieve impressive results, but only referring to the specific context in which they have been trained. As software programmers know very well, trained machines can be “exquisitely well suited to their environment—and ill adapted to any other.”⁵⁴ For example, an algorithm that has to answer a question about drapes in a picture does not look for windows but starts its search from the bottom and stops if it finds a bed (because in the data set used to train it, drapes are found on bedroom windows). The results can be very appropriate for that specific data set, yet do not rely on a knowledge of drapes that can be used in different contexts (e.g., in a classification of fabrics). In fact, the algorithm does not know drapes at all. If general results are needed, one has to reconstruct the group inductively, analyzing many different contexts and aggregating them a posteriori⁵⁵—a procedure that is exactly the opposite of the one of classical statistical science.

MANUFACTURING THE PREDICTED FUTURE

According to the criteria of statistics and modern science, the approach of machine learning presents some fundamental liabilities. Like divinatory techniques, algorithmic procedures are contextual, individual, concrete, and basically obscure. These very aspects, however, are the grounds of their predictive effectiveness. Precisely because they address individual cases and specific contexts, algorithms are expected to predict the future. What is this claim about? And does the forecasting really work?

In machine learning, the predictive ability of algorithms depends on the same factors that make their procedures often incomprehensible to the human mind. Machine-learning algorithms are able to identify patterns in the data that cannot be grasped by reasoning because they are not based on meaning. For the same reason, they cannot be captured by standard statistical procedures that depend on models and data samples artificially selected for some reason. These patterns, however, are expected to disclose the structure of the future regardless of subjects' knowledge and intentions.⁵⁶ Algorithms should find patterns in the mass of unselected observation data, independent from a model.

The lack of a model should lead to a more direct relationship with reality. But what does reality mean when talking about data? The meaning of "real" is very peculiar and refers only to the lack of sampling, that is, to data independence from an interpretive model. This does not imply that algorithms work with "raw data" that come directly from the world. Setting aside philosophical discussions of interpreting reality and the possibility of knowing it, the idea of raw data is very criticizable and has been thoroughly criticized.⁵⁷ Even when the system processes all observational data, the data set on which algorithms work always depends on human intervention: the set includes only the data it includes, could be a different set if it were approached in another way, and has many data points that arise from the behavior and the decisions of people, including decisions about which data is worth collecting in the first place. The procedures of algorithms, moreover, are obviously the result of human design, whether or not the designers themselves know the details of how the machines work. In speaking of "real" data in reference to algorithms,

then, you cannot speak of human neutrality or “rawness” of data due to a lack of human intervention.

That algorithms work with real data does not mean that their data faithfully correspond to the outside world in the sense of classical metaphysics.⁵⁸ Algorithms are not neutral observers who objectively know the world as it is. Algorithms do not know the world at all—“know” nothing. The point is rather that algorithms are themselves real and part of the world in which they operate—from within, not from the outside referring to a model. This changes the meaning of “prediction.” When algorithms make predictions, they do not see in advance an independent external given, the future that is not yet there. This would be impossible. Algorithms “manufacture” with their operations the future they anticipate.⁵⁹ Algorithms predict the future shaped by their prediction.

Predictions are individual and contextual, and refer only to the specific item they address. The algorithms used in predictive shopping, for example, do not reveal how consumer buying trends will be in the next season or which products will have an increased or lowered market share. Instead, algorithms anticipate and suggest which specific products an individual consumer will be willing to buy, even before the individual chooses them, and possibly before someone become aware of a need.⁶⁰ The products can also be ones that the person does not know, but that the algorithm identifies as compatible with their features and with the past choices that they or other similar people accomplished, according to often inscrutable criteria. If the prediction of the algorithm is correct and the person buys the product, this is not because the algorithm saw the future in advance, in part because that future would not exist without this intervention.⁶¹ The person would not have thought to buy that product and may not have even known of its existence. By suggesting the product to the future buyer, the algorithm produces the future and thereby confirms itself—or learns from experience if the suggestion is rejected.⁶² Both errors and correct predictions are useful and help the algorithm to learn, confirming its structures or the need to modify them to take into account new data. In this way, the algorithm becomes increasingly effective in dealing with a world that remains unknown. The same should happen in other cases, such as crime prevention: the prediction should make it possible to act before an individual at risk begins a criminal career.⁶³

The claim that the data processed by algorithms are “real” does not refer to an independent world to be described as accurately as possible, but to the result of a process of “active inference” in which prediction error is reduced “using the twin strategies of altering predictions to fit the world, and altering the world to fit the predictions.”⁶⁴ The world changes as a consequence of algorithms, and algorithms learn from the world how to modify their predictions. Programmers state that “the goal is no longer truth, but performativity. . . . We do no longer (only) decide based on what we know; we know based on the decisions we have to make”;⁶⁵ “expectations are simultaneously descriptive and prescriptive in nature.”⁶⁶

WHEN CORRECT PREDICTIONS ARE WRONG

Despite their limitations, then, algorithmic predictions should always be effective. Even when their anticipations are not realized, algorithms should offer the best possible predictions given the available data, and even the failure of prediction, when it happens, should contribute to learning and improving future performance. If no abstract and general forecasts (as in statistics) are required, for specific cases and in local contexts algorithms should provide accurate and reliable predictive scores, optimizing resource use and enabling humans to detect new possibilities. Tourists discover destinations they would never have thought about and manage to better organize their travels; law enforcement or security agency become more effective; sellers focus their promotions on the relevant portion of a population, avoiding waste and unnecessary annoyance; banks and credit card companies detect more reliable clients and focus their financing on them.

This is not always the case. Critics observe that in some cases the use of algorithms to predict the future may be damaging even when in some sense their predictions are accurate. Harcourt, for example, argues that the increased use of algorithmic tools in criminal law to identify who to search and punish, and how to administer penal sanctions, not only can be morally and politically criticized, but risks undermining the primary goal of law enforcement—namely reducing crime rather than merely increasing arrests.⁶⁷ Reliance on prediction can increase the overall amount of crime,

not only because the increase in attention on specific targets leads to the discovery of “nuisance crimes” that would otherwise go unnoticed and unpursued, or because an initial bias in the data tends to be reproduced by the use of the model, but also because the target population reacts to the targeting effort.⁶⁸ Algorithms are “tools for behavioral modification” whose use must be tempered because they confirm their findings based on the reality they create.⁶⁹

This is the dark side of the performativity of prediction, which reproduces a well-known circularity of divination procedures. If there was no specific intervention,⁷⁰ divinatory predictions tended to be self-fulfilling. As the case of Oedipus shows, everything he did to avoid the predicted outcome contributes to the announced conclusion: he will kill his father and sleep with his mother. In the ancient world the circularity of prediction was regarded as the confirmation of the existence of a higher cosmic order and the negation of chaos. In modern cultures referring to an open future, instead, this circularity results in feedback loops and a serious inability to learn. Algorithms see the reality that results from their intervention and do not learn from what they cannot see because it has been canceled by the consequences of their work. The use of algorithms produces a second-order blindness.⁷¹

The community of programmers is keenly aware of the problems produced by environments that can be changed by the use of algorithms.⁷² The environment can even be actively adversarial, as in the use of algorithms for credit rating purposes that drive people to meet the criteria to which algorithms are oriented. In most of these cases, though, the problems are social rather than technical. Algorithms participate in communication, and this has consequences. According to Harcourt’s argument, for example, if algorithmically profiled persons are less responsive to changes in police policies than others, concentrating crime prevention measures on profiled people can be counterproductive because profiled individuals often have little choice and commit crimes anyway.⁷³ Other areas of the population where surveillance and prevention could be effective instead remain uncovered, and overall crime increases. The algorithm is trained on the world as it was before the action of the algorithm, and thus finds out the most relevant cases: individuals at risk of crime, or products that the user is more likely to buy.⁷⁴ The algorithm then gets

answers about these items, and learns if its predictions were correct or not. But it can happen that the products that the user actually decides to buy, or the individuals who actually commit crimes, continue to escape the prediction because they initially had a very low probability of being targeted, and become more relevant only as a result of the action of the algorithm.⁷⁵ Crime increases because surveillance has moved elsewhere, or niche products become more attractive as a reaction to personalized advertisements of mass products.⁷⁶

The problem is not so much that the punctual prediction of algorithms can be wrong, but that how the future was prepared for is wrong. It is a social problem that must be faced by studying social structures—that is, the environment of prediction. Precise predictions activate reactions that can lead to self-fulfilling or self-defeating circularities, and also (at the same time) to *preemptive* policies, which limit future possibilities.⁷⁷ If decisions are taken today on security measures about individuals who are profiled as possible criminals, their behavior is constrained, but also the options of the decision-maker are limited. If then the crimes turn out to happen somewhere else, one will be watching the wrong people. Or, as in the case of recommendation systems, the use of self-learning algorithms may produce a biased and incomplete view of the future preferences of the users, because the system only sees the responses of users to the recommended items and not to other items, and still doesn't know how they would have reacted to the ignored ones. An algorithm doesn't get any information on users for whom no recommendation has been made, while at the same time targeting mostly clients who were already interested.⁷⁸ The problem in this case is not just the risk of a wrong prediction, but the reduction of future possibilities for all involved actors.

BLINDNESS AND OVERFITTING

The difficulties of algorithmic prediction are different from the ones of statistical forecasting.⁷⁹ The problems do not arise from sampling problems, data shortages, or from the use of misleading interpretive models. Algorithms do not have these worries. Their difficulties depend instead on specific problems of machine learning and in particular on the way algorithms address the relationship between the past and the future.

Algorithms are trained by maximizing their performance on some set of training data, which came from the past.⁸⁰ The predictive effectiveness of algorithms, however, depends on their ability to perform well on previously unseen data that will appear in the future. Training data and real data are as different as the past is different from the future, yet algorithms only know the training data, and the difference between the two sets gives rise to a number of difficulties,⁸¹ which we often are not equipped to face.⁸²

The first consequence is a recognized major problem in machine learning and practice: the problem of *generalization*.⁸³ In machine learning, to effectively generalize means to use what is known to make a prediction about something the algorithm hasn't seen before, as practical experience constantly requires us to do. Every communication, every sentence, every viewing of an object is different from any previous communication or viewing.⁸⁴ How can algorithms deal with the difference between the training data they know and an unknown variety of future data?

Again, the problem is well known to the machine-learning community and intensely debated. Learning algorithms must find a balance between two partially incompatible objectives. Training error must be minimized and the algorithm must learn to process successfully the examples on which it is trained. If not, the problem is *underfitting*: the algorithm has poor performances and is unable to solve complex problems. At the same time, test error should be minimized, increasing the effectiveness of dealing with examples never seen before. If the algorithm learns to work well on the examples given to it, but becomes rigid with respect to each variation, *predictive* error will increase: the algorithm has learned the training examples so well that it becomes blind to every new item. The problem in this case is *overfitting*, which has been called “the bug-bear of machine learning.”⁸⁵ Overfitting arises when the system builds its own rigid and somewhat autistic image of objects, losing the ability to capture the empirical variety of the world. The system is overly adapted to the examples it knows. For example, it learned so well to interact with the right-handed users it has been trained with that it does not recognize a left-handed person as a possible user. In technical terms, the system fails to effectively distinguish relevant information (a signal) from the

irrelevant (noise). In sociological terms, the experience of the past risks undermining the openness of the future.

In conditions of high complexity and high uncertainty, the risk of overfitting increases because the noise component tends to increase more than the signal component—the future tends to become more and more different from the past. There are more elements of the past that should be neglected to effectively predict the future, otherwise the predictions of the system only reproduce the past and its idiosyncrasies. The problem is determining which elements to ignore, that is, to effectively forget. As argued in chapter 5, however, deciding to forget is always a tricky issue.⁸⁶ Overfitting is a risk for all learning systems, especially when learning is performed too long or training examples are rare (few items are observed and in too much detail); however, it is particularly a risk dealing with big data. In very large datasets, often the data are high dimensional and many elements are new.⁸⁷ Elements can be images, handwritten digits, or informal conversations that involve a large number of aspects, many of which are idiosyncratic and different each time. Diversity is so high that even with a lot of available data, the number of examples is still insufficient for the dimensions involved.⁸⁸ In practice it is as if training were always too long and the sample always too small. Learning this past data is not enough to predict the future that does not yet exist.

MEMORY AND FANTASY

About the future algorithms produce, they are and remain blind to it. How can this condition be addressed? In order to avoid the risks of overfitting and corresponding hallucinatory results, machine-learning programmers are often recommended to favor simpler systems, because complexity would tend to increase noise rather than prediction accuracy.⁸⁹ The problem is discussed in terms of the relationship between *bias* and *variance*,⁹⁰ wherein bias measures how accurate the model is, and variance, how different its predictions are from each other. From the sociological perspective bias corresponds to memory and variance to fantasy, with more complex systems tending to have high bias (that is, adherence to the past). That bias is not necessarily wrong (many stereotypes have a realistic

basis), yet is unhelpful on another level because it narrows the focus and prevents models from seeing what does not match their preconceived ideas. Simpler systems would be less accurate but more open, and therefore, more capable of dealing with the unpredictability of the new.

The problem is related to the management of the relationship between the past and the future, which is traditionally the task of memory. An overfitted system practically memorizes the past and uses this skill to predict the future. Hindsight is more accurate than foresight, and the system risks having poor predictive performances and being prone to generalization errors. To deal with overfitting, then, data scientists propose to “drag down” the algorithm by imposing random errors that prevent learning from becoming too accurate.⁹¹ Other solutions recommend forgetting the past altogether: memory is thus seen as a form of bias, and prediction jeopardizes the openness to novelty.⁹²

Is this necessarily the best solution? To better face the future, should one remember worse? As Nietzsche claimed, the ability to forget is crucial to being able to deal with the world, to hope and to learn;⁹³ but on the other hand, without memories, you could not plan and would have to start anew every time.⁹⁴ The result is not better. Those who do not remember the past are not necessarily innovative, and often tend to unknowingly produce trivial and old forms.⁹⁵ The evolution of social memory shows that it is possible to simultaneously increase knowledge of the past and openness to the future.⁹⁶ In modern society the systematic study of the past led to development of the sense of history, but also to the ability to see the future as an open horizon.⁹⁷ The modern age has abandoned the idea that the future reproduces the past and has started to deal with and even value novelty and surprise. Modern society has developed a memory that is capable of forgetting more because it can remember enormously more.

CONCLUSION

The challenge of prediction in our digital society is to combine individual algorithmic forecasting with the openness of the future—a challenge that seems to take the form of the paradox of combining prediction with unpredictability. Can we know our future in advance and still be

surprised? Can we proactively act on coming events without constraining future creativity and the possibility to innovate?

Algorithmic predictions, like ancient divinatory predictions, are contextual, individualized, and basically obscure. Despite the many analogies, however, algorithmic prediction is fundamentally different from divination. Our contemporary world is not the structured universe of divination, in which it was assumed that a global higher order coordinated all phenomena. Even when algorithmic prediction contributes to producing specific predicted events, digital forecast acts in an incomparably more complex, reactive, and unstable social environment than divination—and in a world that does not necessarily have a fundamental order. In divinatory semantics the idea of seeing in advance the structures of the future could be plausible. But in modern society, and even more so in the digitized society in which algorithms work, the intensity of communication is such that any prediction, even if correct, is anticipated, commented on, and reworked, producing new unpredictable complexity with no guarantee of a basic order. An adequate analysis of the effectiveness and the problems of predictive algorithms cannot only be technical, but requires considering the social and communicative conditions of their use.

This is a section of [doi:10.7551/mitpress/14189.001.0001](https://doi.org/10.7551/mitpress/14189.001.0001)

Artificial Communication

How Algorithms Produce Social Intelligence

By: Elena Esposito

Citation:

Artificial Communication: How Algorithms Produce Social Intelligence

By: Elena Esposito

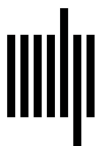
DOI: [10.7551/mitpress/14189.001.0001](https://doi.org/10.7551/mitpress/14189.001.0001)

ISBN (electronic): 9780262368865

Publisher: The MIT Press

Published: 2022

The open access edition of this book was made possible by generous funding and support from the MIT Libraries.



The MIT Press

© 2022 Massachusetts Institute of Technology

This work is subject to a Creative Commons CC-BY-NC-ND license.



Subject to such license, all rights are reserved.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Avenir by Jen Jackowitz. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Names: Esposito, Elena, author.

Title: Artificial communication : how algorithms produce social intelligence / Elena Esposito.

Description: Cambridge, Massachusetts : The MIT Press, [2022] | Series: Strong ideas series | Includes bibliographical references and index.

Identifiers: LCCN 2021013271 | ISBN 9780262046664 (hardcover)

Subjects: LCSH: Telecommunication—Social aspects. | Artificial intelligence—Social aspects. | Online identities. | Social intelligence.

Classification: LCC HM851 .E765 2022 | DDC 303.48/33—dc23

LC record available at <https://lcn.loc.gov/2021013271>

10 9 8 7 6 5 4 3 2 1