

Book Review: Power and Prediction: The Disruptive Economics of Artificial Intelligence (by: Ajay Agrawal, Joshua Gans, Avi Goldfarb)

Szabados Levente^{1, 2}

¹Frankfurt School of Finance and Management

²University of Sopron

L.Szabados@int.fs.de

ORCID: 0009-0003-6046-9361

1 Importance and broader context:

Judging by not just the general news cycle, but the amount of academic publications in the area, the AI revolution is in full swing, and debate is raging about the economic impact of AI, where the potential outcomes are not at all obvious. A general theory about the economics of AI is still being developed, but looking at the literature, it is clear, that it must be in connection with general theories of technological innovation (from economics perspective). In this area, the work of the authors is paramount.

2 Narrower context: Prediction machines

To be able to make sense of this work's contribution, it is essential to take into account the previous, fundamental book from the same authors from 2018, since "Power and prediction" (Agrawal et al., 2022) can be considered a kind of reflection, and further development of the theories laid down in the previous work titled "Prediction machines: The simple economics of Artificial Intelligence" (Agrawal et al., 2018). None the less this is not just a mere "second edition" to the previous, there are substantial modifications, in case some strong revisions of the core theory. With that in mind, the short summary of "Prediction machines" as a foundational piece must be done to give context.

3 Theory of Prediction Machines

The authors tried to interpret AI as a "general purpose technology", much in line with the broader literature, where a general purpose technology is not constrained to be applied in a single economic domain, but can have applications across all possible fields. A good - and often used - analogy in the broader literature is electricity (or the steam engine), which had profound and general impact, reshaping all industrial activity. This viewpoint ties the author's theories strongly to the broader literature of economic innovation (and as we will see, that of "creative destruction" (Ulgen, 2017) or "disruption" (Agrawal et al., 2022))

The most concise way to look at AI - according to Prediction Machines - is, that it is "cheap, good quality prediction", that is, the capability to utilize computers to learn and obtain scalably and cheaply the ability to get predictions based on data about phenomena of interest. For example, the capability to obtain a model trained on previously annotated images of tumour X-ray images, and to use it to produce a probabilistic prediction for a new, yet unseen patient's X-ray. Or, to build a model on some time series data about customer demand for a given article, and use it to forecast future demand etc.

The argument of the authors was: if a given commodity (in this case prediction) becomes cheap, it's consumption increases and it is starting to be utilized as substitute for other goods and services. We use more prediction.

Second corollary to this is, that we start to use prediction as a substitute for other goods and services (since it's cost is low and quality is good), so we substitute for example car driving as an activity with a predictive system for steering decisions, that we call self-driving in vehicles. We substitute things with prediction.

If we can rely on good quality predictions, it is worth changing the business processes accordingly. The example they proposed: if we have good enough pre-diction for what people will buy, we can start shipping before they order. (In fact, this became reality with a patent of Amazon about "Anticipatory Shipping" (Kusnitz, 2014), as well as the case of the company Stichfix's "Ship-then-shop" (Choi et al., 2024) model, that the authors also mention in the later book.) Short: if prediction is good enough, business processes and models can get radically more efficient / overhauled ("disrupted").

Since prediction is based on good quality data, that is data is "the new oil" (Talagala, 2022) for the "engines" of prediction, it's value will go up and become a strategic asset for organizations

Maybe most importantly, the authors introduce the distinction between "prediction" and "judgement" in a decision. Former (e.g. "this is a tumour with 75 percent probability") can be of great service for the latter ("it is worth it to immediately do an exploratory surgery, even when it is risky for the patient"), but the second is uniquely human and valuable. Gaining information about the future (prediction) is not enough, since decisions entail value judgement and analysis / taking of risk. This means, that the value of human judgement - considered to be a "complementary good" for the commodity of prediction goes up.

This latter will be the key, the relationship between prediction and judgement in the insights of the next book of the authors.

4 Power and prediction: going beyond the basic theory

As an honest and laudable starting point: the authors admit that they were wrong. They predicted the rapid spread of AI and profound transformation of nearly all business domains. They were wrong, in the sense that AI had great successes in given domains, and is obviously on the agenda of nearly all companies, but with all successes, it was more deployed in narrow niches, and did not usher in the great wave of change they predicted.

To elaborate on this, and in strong connection to a broader set of concepts about the "lifecycle of innovation" (David, 1989), the authors introduce three concepts of technology application: point solutions, applications and system solutions.

Point solutions can be considered technologies that are replacing a single step in the current business process, making it more efficient without any modification to the process itself (doing the same drilling at the same spot of the factory, but instead of steam, with an electric drill), while applications are local solutions where the reorganization of single processes become feasible (a whole drilling / milling station electrified, doing things differently), though keeping the general system unchanged. System changes (like reorganizing the whole layout and processes of the factory to accommodate electric equipment all around) are profound changes of how we do things, having the greatest potential impact - albeit with the greatest investment.

With this in mind, it has to be noted, though, that the authors were not wrong in the sense that they misjudged technology or the main economic drivers, but in the sense of not considering other important factors, and were surprised to see, that point solutions are popping up everywhere, applications are rarer, system change is largely absent yet with respect to AI. In a sense, their second book is dedicated to elaborate, why.

5 More detailed relationship between prediction and judgement

After introducing the dilemma of point solutions vs. system change, the authors recur to elaborate the concept of judgement and prediction more in detail. Their argument is, that in all cases of decisions, the two elements are necessarily present, albeit two more observations are made:

Decision is "costly", in the sense of cognitive load and responsibility (which by the way ties in well with the concept of "decision fatigue" (Tierney et al., 2011) developed in the psychology literature, partly also quoted by the authors), thus human beings must necessarily "economize" the usage of their judgement as a precious, constrained resource. (We can take in only so much information, and form only a handful of well thought out value judgements about what should be done as an action.) Because of this, there is a natural tendency to forego decisions, especially in the form of fixed rules that are helping to avoid decision situations but introducing hidden uncertainty, as well as rigidity into organizations' behaviour.

On the other hand, the usage of rules is a key contributing factor for stability, since it allows for the efficient coordination of activities across an organization by foregoing the necessity for constant communication, reducing the coordination overhead.

In this sense, AI technologies - so prediction - is basically reintroducing the necessity for applying judgement, thus on the one hand causing extra cognitive load, as well as making coordination more complicated (by challenging the rules), but also holds a great promise, if "well oiled", so flexible processes can dominate the business, or if decision processes can be dynamically adjusted.

The separation of the aspects of "prediction" and "judgement", which were practiced together by humans previously is the second major effect the authors study. They intend to show, that this new division is even more challenging, since it shifts decision power from the people, organizational

units or groups that had control over them previously. As they demonstrate it with several community and company examples, the expertise needed for jointly doing prediction and judgement exercised by experts is now "factored" into scalable machine prediction, and either a new set of experts specializing in judging the probabilistic predictions of AI models, or to committees that set up decision criteria (basically pre-judgements of value). In this latter case, judgement can also be scalable, but requires a complete overhaul of the responsibility and decision system.

It is worth noting, that in this light, much of the literature on organizational change, social epistemology and decision making comes into play, so this represents a deep connection point to other literature that would definitely be worth exploring more in detail.

6 Evaluation of the book's contributions

With that in mind, let us now move to a more critical evaluation of the contributions of the book.

On the positive side, it is worth strongly emphasizing, that the book is filling a large gap in the understanding of AI's economic potential, and is especially useful for those practitioners that want to use it as a kind of thought framework for implementing AI solutions into organizations (this is a clearly intended audience of the book, demonstrated by the introduction of such conceptual aids as the "AI Systems Discovery Canvas" and so on.)

On the exceptionally positive side it is also clear, that the book is invaluable in providing a thoroughly thought-out model that can guide the inquiries into the economic impact of AI, also connecting it to the broader field of organizational behaviour, and behavioural economics in general.

On the criticism side - maybe not completely fairly - one can mention the fact, that due to some recent events, namely the introduction, and subsequent rapid mass adoption of Large Language Model based solutions (notably OpenAI's Chat- GPT and alike), some of the insights of the book became nearly instantly a little dated.

To be more specific: the book does not concern itself with the possible broad (and currently rapidly advancing) adoption of AI tools in the "business to customer" domain, so the whole slew of changes (in personal productivity, consumption of information, cognitive labour etc.) that are coming from the

individual's side, and enabling new business models and huge competitive pressure on the labour market. One could argue, that this is just one (or a few) company's success with business model innovation enabled by AI, which is true, but gives less credit to these phenomena than they deserve.

Going forward: the strong push towards more generally applicable AI models ("foundational models" (Bommasani et al., 2022)) would merit some new considerations. In many of the expositions of the book, AI solutions, even when system like, are focusing on (from the technological sense) narrow, domain specific AI capabilities, thus positing competitive advantage in a single domain for a well "oiled" "feedback flywheel". In practice, though, the training and deployment (and often even Open Source release) of generally capable AI solutions and agents (Xi et al., 2023; Durante et al., 2024) are questioning the defensibility of solutions, systems and companies that have a narrow AI solution, in stead foreshadowing, that the real advantage may lie in the hands of "AI first" organizations, that come from purely the domain of AI research and innovation, and intend to disrupt many business areas simultaneously. This is a kind of phenomena that the authors simply did not have the opportunity yet to study.

Another - maybe less severe - challenge towards the book's theories comes from the apparent ability and usage of current state-of-the-art Large Language Model based systems in planning and task decomposition (Kojima et al., 2023), in which they are autonomously decomposing general goals given by humans into actionable tasks and carry out their execution with a broad sense of "self control". One could argue, that this case does not violate the premise of the book about human judgement, since still, the setting of goals and the evaluation of the final results is still in human hands, but none the less, this phenomena deeply alters the meaning, or the type of judgement that is necessary in these cases.

Also, it is worth noting that throughout the book, the topic of execution automation is not reflected upon clearly. If we take such passages as: "AI provides predictions, and hence, value is created by improved decision-making. Point solutions allow prediction to improve existing decisions." We can feel, that there is another aspect which is not sufficiently emphasized: much of automation, that is, automatic execution has a non-negligible cognitive, hence decision element in it, (see (Varela et al., 1991)) thus: good quality prediction is not just an input for decisions, but also a vital prerequisite of automated execution! The real breakthrough might come

when automated execution of tasks (sometimes in a digital, sometimes in an "embodied" physical modality) is perfected!

The complex problems and necessary skills with the widespread adoption of human value judgement based automated execution are still to be investigated, and loom large on the near horizon.

7 Takeaways and recommendation

The main dilemma of the book is extremely timely, the contributions of it in build- ing the thinking framework invaluable.

The question is, from the business practice:

"Large companies rarely find it worth it to transform the way their industry operates, especially if their industry is currently profitable. The risk of getting it wrong is too high.", so while "system change" and "disruption" is the promise (and also the challenge), maybe incumbent organizations are at an inherent disadvantage...

If the hypothesis is true, that economic power does not disappear, just shifts, where will it shift to? Towards the organizations having the largest models? With AI training know how? With application know how? With data? With capital? With judgement surplus? With ones that build new economic / governance systems for old business problems? ("However, those who do not have a stake in the current system are often best poised to reap the reward from creating a new one.")

And where does the economic power of labour shift?

"As Christensen saw it, the problem was the velocity of history" - how can organizations speed up? By automation / computerization / agentification?

But how can individuals speed up??? Mental training? Using AI to speed up individual learning? With human self augmentation?

As the book argues: "Automation requires codifying judgment. A human must specify judgment when the machine is deployed, rather than upon receiving a prediction. That means that judgment has to be useful for a large number of decisions, and it has to be described in a way that can be coded. That isn't so easy." So maybe abstract, structured thinking is the key trait for human employability?

If the fact that the amount of judgement (if scales in pre-judgement) needed to operate our business processes require way smaller number of way higher qualified specialists than the number of people working now. Comparatively speaking judgement will take on a small amount of working time, thus, at current demand levels, it is not needed in such quantities to mandate mass employment.

Finding answers to these questions, and empirically validate the claims of the theoretical framework proposed is thus an interesting area to work on.

References

Agrawal, A., Gans, J., & Goldfarb, A. (2022). *Power and prediction: The disruptive economics of artificial intelligence*. Harvard Business Review Press.

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines*. Harvard Business Review Press.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., Donahue, C., Doumbouya, M., Durmus, E., Ermon, S., Etchemendy, J., Ethayarajh, K., Fei-Fei, L., Finn, C., Gale, T., Gillespie, L., Goel, K., Goodman, N., Grossman, S., Guha, N., Hashimoto, T., Henderson, P., Hewitt, J., Ho, D. E., Hong, J., Hsu, K., Huang, J., Icard, T., Jain, S., Jurafsky, D., Kalluri, P., Karamcheti, S., Keeling, G., Khani, F., Khattab, O., Koh, P. W., Krass, M., Krishna, R., Kuditipudi, R., Kumar, A., Ladhak, F., Lee, M., Lee, T., Leskovec, J., Levent, I., Li, X. L., Li, X., Ma, T., Malik, A., Manning, C. D., Mirchandani, S., Mitchell, E., Munyikwa, Z., Nair, S., Narayan, A., Narayanan, D., Newman, B., Nie, A., Niebles, J. C., Nilforoshan, H., Nyarko, J., Ogut, G., Orr, L., Papadimitriou, I., Park, J. S., Piech, C., Portelance, E., Potts, C., Raghunathan, A., Reich, R., Ren, H., Rong, F., Roohani, Y., Ruiz, C., Ryan, J., Ré, C., Sadigh, D., Sagawa, S., Santhanam, K., Shih, A., Srinivasan, K., Tamkin, A., Taori, R., Thomas, A.

W., Tramer, F., Wang, R. E., Wang, W., Wu, B., Wu, J., Wu, Y., Xie, S. M., Yasunaga, M., You, J., Zaharia, M., Zhang, M., Zhang, T., Zhang, X., Zhang, Y., Zheng, L., Zhou, K., & Liang, P. (2022). On the opportunities and risks of foundation models.

Choi, W. J., Liu, Q., & Shin, J. (2024). Predictive analytics and ship-then-shop subscription. *Management Science*, 70(2), 1012–1028.

David, P. (1989). Computer and dynamo: The modern productivity paradox in a not-too distant mirror. *The Warwick Economics Research Paper Series (TWERPS)*, University of Warwick, Department of Economics.

Durante, Z., Huang, Q., Wake, N., Gong, R., Park, J. S., Sarkar, B., Taori, R., Noda, Y., Terzopoulos, D., Choi, Y., Ikeuchi, K., Vo, H., Fei-Fei, L., & Gao, J. (2024). Agent AI: Surveying the horizons of multimodal interaction.

Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., & Iwasawa, Y. (2023). Large language models are zero-shot reasoners.

Kusnitz, S. (2014). Why Amazon’s anticipatory shipping is pure genius. *Forbes*.

Talagala, N. (2022, March). Data as the new oil is not enough: Four principles for avoiding data fires. *Forbes*.

Tierney, J. (2011, August). Do you suffer from decision fatigue? *New York Times Magazine*.

Ulgen, F. (2017). Creative destruction. In *Springer New York* (pp. 1–8). Springer.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press.

Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., Zhang, M., Wang, J., Jin, S., Zhou, E., Zheng, R., Fan, X., Wang, X., Xiong, L., Zhou, Y., Wang,

W., Jiang, C., Zou, Y., Liu, X., Yin, Z., Dou, S., Weng, R., Cheng, W., Zhang, Q., Qin, W., Zheng, Y., Qiu, X., Huang, X., & Gui, T. (2023). The rise and potential of large language model based agents: A survey.