A conversation with Alexander Gammerman

Zhiyuan Luo

Department of Computer Science,
Royal Holloway University of London, TW20 0EX, UK
Zhiyuan.Luo@rhul.ac.uk
https://cml.rhul.ac.uk/people/zhiyuan/

Abstract. Alexander Gammerman is a British computer scientist and Professor at Royal Holloway University of London. He is the co-inventor of conformal prediction. Alex is the founding director of the Centre for Reliable Machine Learning at Royal Holloway, and a Fellow of the Royal Statistical Society. He has published 9 books and about 200 research papers. He ranks amongst the top 1% researchers in Artificial Intelligence and Machine Learning.

Alex's work spans two countries (the Soviet Union and the United Kingdom) and two centuries: the second part of the 20th century and the first quarter of the 21st century. It has been an extremely interesting time with the rapid development of Artificial Intelligence, Machine Learning, the Internet and the Large Language Models. How are these discoveries, inventions, and developments influencing the life of an academic?

I have known Alex personally for many years. I first came as a PhD student to Heriot-Watt University, Edinburgh, in 1988. I then worked in various places before arriving at Royal Holloway University of London as a Lecturer in 2002. Alex has been more than a mentor and colleague to me — he has been a guiding figure and a close friend, with his pioneering work continuing to inspire both me and my PhD students.

The following conversation, which took place both face-to-face and remotely in July–August 2025, offers a glimpse into Alex's remarkable scientific journey, his reflections on decades of research, and his perspective on the ever-evolving world of Artificial Intelligence.

1 Family

ZL: Alex, thank you for agreeing to have this conversation. I am sure you have many interesting stories to tell. To begin with, could you tell us a little bit about your family background?

AG: I was born in Alma-Ata (USSR, Kazakhstan) in 1944, where my parents were evacuated in 1941, when the Germans approached Moscow in the 2nd World War. After the war, the family returned to Moscow and soon after, my dad, an expert in the textile industry, was invited to Tallinn to help with the development of the light industry in Estonia; so the family moved to Tallinn. That's where I grew up and finished school.

My parents were born in Ukraine and met in Odessa at Odessa National Economics University in 1924. Dad graduated from the University, but my mother was expelled since she was the daughter of a cantor and therefore, by the "politically correct" communist ideology, did not have the right to study at any university.

My dad - Yakov Moiseevich Gammerman (1895-1978) - was working all his life in the textile industry eventually became the director of a large textile factory Marat in Tallinn. The factory was famous in the USSR for its high-quality men's and women's underwear. There is a website describing my dad's life (in Ukrainian)[1].

My mum - Susannah Grigoryevna Gammerman, (née Elisheva Darer, 1906-1996) - stayed at home while my brother and I were growing up. After we left home for education and work, she worked as a Librarian at the Library of the Academy of Sciences of Estonia.

Fig. 1. Dad and Mum with my brothers, Mikhail and George, 1939

My parents had 4 children - 4 boys - but two of them died during the 2nd World War, and only my eldest brother Misha and me, the youngest, survived.

My older brother Misha - Mikhail Yakovlevich Gammerman (1931-2006) - was the Chief Designer of flowmeters, the devices to measure the flow of heat and liquid. He received many awards and medals for his inventions. The flowmeters

[1] http://berdychiv.in.ua/гаммерман-яків-мойсейович/

were of very high accuracy and used all over the Soviet Union. You can find more details in a Wikipedia page devoted to him (in Russian)[2].

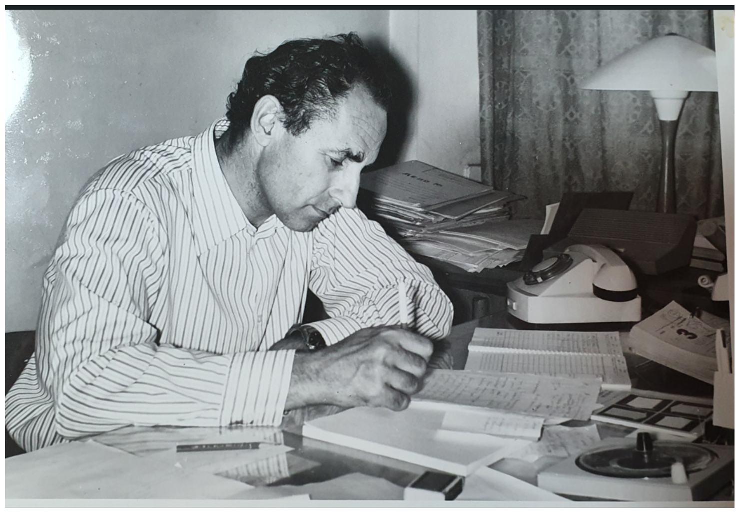

Fig. 2. My brother Misha, 1975

Apart from that, my brother was a great chess player and participated in various tournaments in the USSR. He had a great influence on my upbringing and education.

I met my wife, Suzy, in 1979 in St. Petersburg, while she was a visiting student from Cambridge University. We married in 1982 and emigrated to the UK in 1983. Our 3 children were born and educated in Britain: James (b.1990), graduated from Imperial College London, specialist in Data Science; Anya (b.1992), graduated from King's College London in neurosciences; Sonia (b.1996), graduated from University College Maastricht, specialist in communications and marketing, but her real passion is electro swing jazz music and singing.

[2] https://ru.wikipedia.org/wiki/Гаммерман,_Михаил_Яковлевич

Fig. 3. L-R: Anya, Suzy, me, James, Sonia in Trafalgar Square, 2022

2 School, University and Postgraduate Education, 1963-1972, the Soviet Union.

ZL: Your parents lived through the Russian Revolution and the Second World War. That must have shaped their outlook and yours. Their resilience and drive to succeed clearly influenced your academic path. What was school like for you? Any favourite subjects?

AG: School was fine, nothing remarkable. Most lessons were dull, to be honest - except for Chemistry and Literature. We had two genuinely passionate teachers, and that made all the difference. Outside of class, I kept busy with basketball, volleyball, and athletics.

ZL: After school, you chose to study physics at university. What drew you to it? And what was life like at St. Petersburg University back then?

AG: In my family, the cult of education was very high: my father was educated as an economist and then continued learning various aspects of the cotton industry all his life. He is the author of two books devoted to textiles and rose from a foreman to the director of a big factory. My brother Misha graduated from school with a gold medal, graduated from the University (he studied at Moscow Aviation Institute and then at Tallinn University of Technology), created his own laboratory and was an outstanding innovator. My grandparents from my mum's side (and several previous generations) were also well-educated, religious people. So yes, you could say my path was laid out early.

I got hooked on physics in high school after reading Feynman's Lectures. Quantum physics fascinated me - and that curiosity stuck.

Back then, physics was everywhere - in films, books, public discourse. It had a kind of glamour. After the war and the rise of nuclear science, physics felt more vital than art. Studying it was like being part of the future. Like doing AI or machine learning today. I participated in various Olympiads in physics and was even among the winners at some competitions. I also read a very interesting book written by Daniil Danin, called *The Inevitability of a Strange World*. This book is about how the basic ideas of quantum mechanics were developed and who were the people behind it: Bohr, Schrödinger, Heisenberg and others. That sealed the deal.

ZL: And once you got to university - what did you actually learn?

AG: I got a solid grounding in all kinds of physics - mechanics, quantum electrodynamics, relativity. Gell-Mann's quark theory was emerging, but it wasn't part of our curriculum yet. Oddly enough, the subjects I'd need later - probability theory, programming, information theory - were barely covered. I had to pick those up as a graduate student.

And yes, I learned a few extracurricular things too: to drink vodka, play cards in all-night sessions, chase girls ... the usual student fare. [Laughs].

But mostly, I spent time in the library, reading math and physics books. The lectures were hit-or-miss. None of the professors really sparked my imagination. Eventually, I grew disillusioned with physics.

ZL: Still, were there advantages to studying it?

AG: Absolutely. Physics gave me a broad education and a strong foundation in maths and science. That opened doors to all sorts of fields - mathematical biology, computational science, machine learning. I don't think any other discipline would've offered such range. It taught me how to model problems, make generalisations, and trust physical intuition—even when the maths didn't quite behave. Eventually, I started carving out my own path and chose to pursue mathematical biology in postgraduate studies.

2.1 Mathematical Models in Biology; Agrophysical Research Institute, St. Petersburg, 1969-1972.

ZL: Switching from physics to mathematical biology sounds like quite a leap. What prompted that change?

AG: Well, theoretical physics jobs were scarce. I had to choose - either stick with experimental physics or find something new. Laser tech was popular, and many of my classmates jumped on that bandwagon. But I've never liked following the crowd, and pure experimentation didn't appeal to me. For my final-year, I found an interesting project in biophysics—modelling protein structures. It involved using computers for calculations, which nudged me toward programming. That was a turning point.

ZL: So that's when your academic life really began?

AG: Yes, properly so. I enrolled in postgraduate studies at the Agrophysical Research Institute (AFI) in St. Petersburg. There were entrance exams—Physics, English, Philosophy—and in April 1969, I got in as a PhD student. The scholarship was 69 rubles a month, about 17 dollars. Not exactly enough to live on, so I started tutoring physics and maths on the side.

ZL: What was the Institute like?

AG: It turned out to be a great place to study. Founded by Academician Abram Joffe (he also set up the Institute of Physics and Technology in St. Petersburg) it became a hub for Mathematical Biology, and it attracted some brilliant minds. The scientific life was buzzing - seminars, conferences, lively discussions. We had visits from top researchers like Alexey A. Lyapunov (a mathematician), Nicolay V. Timofeev-Resovsky (a biologist), Raisa L. Berg (a geneticist) and others. It was intellectually rich.

Social life was vibrant too - skits, wall newspapers, concerts. But of course, everything was under Party control. Eventually, a new Party boss arrived and started purging the "unreliable" ones - people who were too independent, too free-thinking. It was a shame. Many talented researchers were forced out.

ZL: It sounds like a stimulating place against the political trends in the country. Who supervised your PhD?

AG: Dr. Leonid Fukshansky. He was just a few years older than me and had done impressive work on modelling the biological clock in plants. I read his thesis and was thrilled when he agreed to supervise me. He proposed a project on modelling phytochrome - a plant receptor involved in photomorphogenesis. I had no idea what morphogenesis even meant! But I went to the library and, by sheer luck, stumbled upon a little-known Alan Turing's 1952 paper *The Chemical Basis of Morphogenesis* [24]. That was a good enough motivation and sparked my interest.

ZL: That's quite a discovery. Did Leonid remember things the same way?

AG: Not exactly! [Laughs]. According to him, a colleague asked if he'd take on another PhD student. He was hesitant - already had several. But when he heard the student was a USSR schoolboy athletics champion, he changed his mind.

ZL: You were a champion?

AG: Yes, back in school I did sprints, hurdles, long jump. I was part of Estonia's 4x100 relay team in the 1962 Youth Spartakiad. We won the final. After school, I left athletics behind, but apparently Leonid was still impressed.

ZL: And as a supervisor?

AG: He was excellent. He pointed me toward interesting open problems and useful references. He also taught me how not to be the arrogant mathematician barging into biology with grand theories but little understanding. And equally, how not to be the biologist who ignores mathematical rigour. That balance was crucial.

ZL: Could you explain the *phytochrome* system? How does it work, and why is it important in plant biology?

AG: Phytochrome was discovered, mainly by the work of Harry A. Borthwick and Sterling B. Hendricks in 1959 in Beltswill (USA). The presence of phytochrome allows us to understand how plants control their germination, stem growth, leaf growth, and flowering [2].

Hendricks proposed that the phytochrome exists in two inter-convertible forms, one absorbing in the red (P_r) spectrum at a maximum of 660 nm, and one absorbing in the far red (P_{fr}) spectrum at approximately 720 nm, and concluded that the red and far-red light causes transformation between the two forms - a photoreversible reaction[3].

ZL: Fascinating. Let's talk about your PhD. What were the main questions you tackled, and what do you see as the key contributions of your research?

[3] https://www.degruyter.com/document/doi/10.1515/ci-2016-0506/html?lang=en

Fig. 4. Leonid Fukshansky, c.2020

AG: My work focused on two main problems. First, I wanted to understand the dynamics of phytochrome—a light-sensitive pigment in plants. Second, I aimed to find a link between those dynamics and actual physiological responses, like how plants grow.

To model the dynamics, I had to build a system of first-order differential equations. The tricky part was that the coefficients weren't just numbers—they were functionals that depended on the light spectrum and the reactions happening in the phytochrome system [8,13,14].

Later, I extended the model to handle more complex situations, like when the light spectrum changes over time or varies with depth in the plant tissue. That turned the equations into non-autonomous differential equations, which don't have neat, closed-form solutions. So I turned to numerical simulations and started using computers to solve them.

Most of my time was spent running simulations - changing initial conditions, tweaking light properties, adjusting parameters - to see how the system behaved. That's how we could predict the behaviour of phytochrome under different lighting conditions.

ZL: And the second part - connecting that to plant responses?

AG: Yes, once we had a handle on the dynamics, we looked for correlations with actual plant behaviour. One clear example was stem elongation. We found a measurable relationship between the concentration of the active form of phytochrome and how much the plant grew. That was a big step. It showed that the mathematical model wasn't just abstract—it could explain real biological effects. We published those results in a Soviet journal, and they were later translated into English [3].

ZL: That must have drawn some attention.

AG: It did. Our work was noticed by researchers abroad, in Europe and the USA - people like Professor Hans Mohr in Germany and Professor Harry Smith in the UK. When I eventually came to England, Harry Smith even tried to get me a position in his lab in Leicester. But by then, I was already a lecturer in Edinburgh.

ZL: And your PhD?

AG: I defended my thesis and received my PhD in physics and mathematics (кандидат физико-математических наук). Looking back, I think the real contribution was laying the groundwork for quantitative modelling of phytochrome. Today, of course, we know much more - phytochrome has even been found in microorganisms. But those early models helped open the door.

2.2 Pattern Recognition; Regional Research Computer Centre, St. Petersburg, 1972-1983.

ZL: What came next after your PhD?

AG: My first job was actually just down the corridor - from the same Agrophysical Research Institute, but in a different department: the Information Systems Lab. Later, around 1975-76, it evolved into the Regional Computing Centre of the Agricultural Academy of the USSR.

The Centre was in a beautiful old building right on St. Isaac's Square in the heart of St. Petersburg. It belonged to the All-Union Institute of Plants, or VIR, which housed the World Collection of Plants. Our job was to process and analyse that vast collection.

ZL: That sounds like quite a shift.

AG: It was. The head of the Centre, Dr. Mark Lanin, was a remarkable man - brilliant researcher, charismatic, and, incidentally, a fantastic chess player.

He'd developed two major systems. One was TABIA, an accounting system that came out in the mid-70s - just before the first spreadsheet appeared in the

West. The other was SL-740, a search algorithm that was probably the fastest in the USSR at the time.

ZL: And what was your role?

AG: I was tasked with developing statistical and pattern recognition algorithms. We applied them to medical diagnostics and plant classification. That's when I really started diving into statistics and what we now call machine learning.

Back then, AI wasn't mainstream in the USSR. The buzzword was "Cybernetics" - which had just recovered from being labelled a "bourgeois pseudoscience" in the 1950s. By the 70s, it was all the rage, along with control theory.

ZL: So that's when the shift toward machine learning began?

AG: Exactly. I didn't know it at the time, but a book by Vapnik and Chervonenkis on pattern recognition had just come out [31]. Years later, I'd meet both of them and even work alongside them. But back then, I was just starting to explore the field - without knowing where it would eventually lead.

Fig. 5. Mark Lanin, c.2000

ZL: Could you walk us through the kinds of projects you worked on at the Centre? And what sort of computing tools did you have at your disposal?

AG: Sure. One of our first major projects was in pediatric oncology. We collaborated with Dr. Genrich Fedoreev at the St. Petersburg Oncological Institute—he headed the children's department. It was quite pioneering at the time: using statistical methods and computers to help diagnose and treat cancer.

We focused on hemangiomas, which were usually treated surgically right after birth. Sadly, that often led to permanent cosmetic or functional damage. But by applying pattern recognition techniques, we showed that most of these cases didn't need surgery at all. They tended to regress naturally by the age of five or six [32]. That was a big moment-research saying, "Wait, maybe less is more."

We later extended the same statistical approach to study Hodgkin's and non-Hodgkin lymphoma in children - looking at diagnosis, treatment strategies, and prognosis.

ZL: That sounds like a lot of data to work with.

AG: It was. And most of it was qualitative - categorical features, not numbers. So we had to figure out which features mattered most, and how they were related. We developed a multidimensional contingency coefficient based on the Kullback-Leibler information measure [17]. Quite elegant, actually.

As for computing - well, we had a MINSK-32. It was the Soviet version of the IBM 360. We programmed in Assembler and Fortran, and everything was stored on punched cards. You had to be very careful - drop a box of cards and your entire week's work could scatter across the floor.

ZL: And this was all happening at the Centre?

AG: Yes, the Computer Centre was buzzing with activity. We didn't just work on medical data - we also applied similar techniques to the World Plant Collection housed at VIR. That led to several publications [17,33,16].

The atmosphere at the Centre was wonderful. Lively discussions, visitors from all over the USSR and abroad. It felt like a real intellectual community.

But outside those walls, things were changing. The Afghan war had started. Antisemitism was creeping back. Dissidents were being jailed. The mood was darkening. We were refining predictive algorithms while living in a country where, as we joked, "the past was more unpredictable than the future." [Laughs].

ZL: Was the Centre eventually shut down?

AG: Yes. The Party decided we were too free-spirited, not ideologically reliable enough. Institutions like ours were quietly dismantled. It was a sobering moment. Many of us realised the country no longer had a future we could believe in. In the end, it was clear: in the battle between "Bayesianism and Bolshevism"[4], science didn't stand a chance.

[4] Bolshevism - Political theory and practice of the Bolshevik Party which, under Lenin, came to power during the Russian Revolution of October 1917. The Bolshevik (meaning 'majority') radical communist faction within the Russian Social Democratic Labour party (*Oxford Reference).*

3 Machine Learning, Heriot-Watt University, Edinburgh, Great Britain, 1983-1993.

ZL: You arrived in the UK in 1983. What was that experience like - and how did your first year unfold?

AG: We arrived in early May - Suzy and I, fresh from St. Petersburg. It was a train journey across Europe, and the contrast between East and West was striking. I still remember the border crossing in East Berlin: dim lights, grim-faced soldiers. Then suddenly - West Berlin! Bright, bustling, full of music and life. It felt like stepping into another world.

Once in England, I started applying for jobs. To my surprise, I discovered that some of my papers from Soviet journals had already been translated into English by American publishers. I listed those in my CV, and that helped me land my first job: Lecturer in Computer Science at Heriot-Watt University in Edinburgh, Scotland starting October 1983.

ZL: Why Heriot-Watt?

AG: Funny story. At the interview, they asked why I'd applied. I said, "Well, your university is famous for AI and machine learning - people like Donald Michie work here." The panel looked puzzled. "That's Edinburgh University," they said. I blinked. "But this is Edinburgh University, isn't it?" "No, no - we're Heriot-Watt. Different place."

In the Soviet Union, each big city had just one university. The rest were polytechnic institutes of higher education. So my confusion was understandable - and we all had a good laugh. Somehow, despite that mix-up, they offered me the job.

ZL: What was the department like?

AG: Heriot-Watt was keen to build up its research in AI. The so-called "AI winter" was ending, and Japan's push for fifth-generation computing had reignited interest. Professor Howard Williams, who headed the department, secured funding to launch a new MSc course in Knowledge-Based Systems. Several of us were hired to teach and develop research in that area.

ZL: And you stayed for ten years. What do you remember most?

AG: The department was in Grassmarket, right in the heart of Edinburgh. It was a magical place - medieval gates, lively pubs, the avant-garde Traverse Theatre just around the corner, and even a tiny dog cemetery hidden behind a wall near Heriot School. The 1st of May bookshop was next door, where my friend and colleague Greg Michaelson volunteered. It was a wonderful mix of history, culture, and eccentric charm.

Academically, I kept pursuing machine learning. But at the time, it wasn't fashionable. Everyone was into logical programming and Prolog. I was pretty much a one-man band in machine learning at HWU.

Fig. 6. North Berwick, Scotland, c.1985

I even remember a memo from the AI department at Edinburgh University advising PhD students not to choose machine learning - it was considered too ambitious for a three-year thesis. That made me smile. Sometimes, being out of step with fashion is exactly where you need to be.

ZL: Aside from your research, what courses did you teach at HWU?

AG: The very first course I was asked to teach at Heriot-Watt was in Artificial Intelligence. That sounded exciting - until I tried to find a textbook. This was 1983, and to my surprise, there were simply none to be found. I scoured bookshops in London - Dillon's and several others - and then tried my luck in Edinburgh. Nothing.

So I had to build the course from scratch. I started by writing notes in Russian, then painstakingly translated them into English. After that, I transferred everything onto transparency slides - those old acetate sheets we used with overhead projectors. It was quite a task, but also strangely satisfying. You learn a lot when you have to teach from your own notes.

Most of my teaching revolved around the MSc in Knowledge-Based Systems. I supervised a good number of final dissertations, and in the first year, I also taught a course in Computer Vision.

At the time, Edinburgh was buzzing with AI research. Donald Michie's team at Edinburgh University was doing fascinating work - developing algorithms and testing them on chess-playing programs. The idea was simple: if a computer could beat a human at chess, we'd achieved AI. Looking back, it was a bit naïve, but it captured the imagination.

My brother, as I've mentioned, was a serious chess player. He had an enormous collection of chess literature, and when he sent me some of it, I passed it along to Michie's team. They were thrilled. I became a regular visitor to their lab, and those exchanges - between science, games, and curiosity - were some of the most enjoyable moments of that time.

3.1 Bayesian Inference without assuming independence

ZL: You've done some fascinating work on Bayesian inference at Heriot-Watt. Could you tell us more about those projects?

AG: Once I'd settled in - got my lecture notes in order, learned my way around UNIX and the hardware - I was keen to dive back into research. My focus was still pattern recognition and machine learning.

One thing that struck me early on was the divide in British statistics between the Bayesian and frequentist camps. In Russia, that split wasn't nearly as pronounced. What surprised me even more were papers claiming that probability theory wasn't suitable for AI or reasoning under uncertainty. Some of those arguments were, frankly, quite muddled-confident, but conceptually off.

ZL: So you leaned into Bayesian methods?

AG: Yes, almost inevitably. At Heriot-Watt, I started a project to challenge a peculiar claim floating around—that Bayesian models couldn't be used in knowledge-based systems because they required independence between features. We set out to show that wasn't true.

We used a substantial medical dataset, thanks to Dr. Tony Gunn and Dr. Steve Nixon from Western General Hospital in Edinburgh. Their clinical insights were crucial. We compared two models: Proper Bayes, which made no independence assumptions (affectionately dubbed the G&T system, though not for Gin and Tonic), and Simple Bayes, which did assume independence.

ZL: And you worked with your father-in-law on this?

AG: Yes, Roger Thatcher. He was an extraordinary statistician of remarkable breadth, whose intellectual curiosity ranged from demography to cosmology, machine learning to archaeology, and who, despite never holding an academic post, published widely with great clarity and rigour across disciplines.

One of his most elegant contributions, the 1964 paper "Relationships between Bayesian and confidence limits for predictions" [22], extended the notion of confidence limits to apply to predictions about future observations and not just to unknown parameters - a subtle but profound shift. In doing so, he built a conceptual bridge between Bayesian and frequentist paradigms, showing that under certain conditions, frequentist procedures could be interpreted as Bayesian ones with specific priors.

It was a beautiful insight, especially in prediction problems. At a time when the two camps were seen as fundamentally opposed, Roger's work quietly demonstrated that they weren't so different after all.

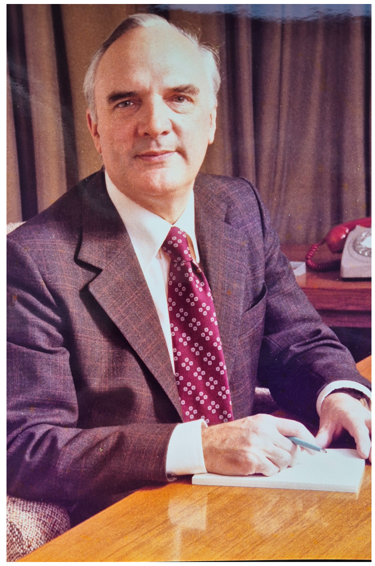

Fig. 7. Roger Thatcher, 1980

The project, indeed, demonstrated that the assumption of independence is not required. The results were summarised in our joint paper "Bayesian diagnostic probabilities without assuming independence of symptoms" [9].

ZL: When I joined Heriot-Watt as a PhD student in 1988, you were working on Bayesian Belief Networks. Could you tell us a bit about that line of research?

AG: Yes, that was an exciting time. Bayesian Belief Networks - or BBNs - were a kind of middle ground between the Simple Bayes and Proper Bayes models. The idea was that not all features had to be independent - just some of them.

If you imagine a graph where the nodes represent variables and the edges represent dependencies, then any two nodes not connected are assumed to be conditionally independent. Add a set of conditional probabilities to that graph, and voilà - you have a Bayesian Belief Network. It's a directed acyclic graph, meaning no loops, and it gives you a structured way to reason under uncertainty.

Judea Pearl had just developed his message-passing algorithm in 1982 [18], which made inference on these networks computationally feasible. That really opened the door for practical applications.

ZL: And that led to the Offender Profiling project?

AG: Exactly. It was called the *CATCHEM* project - a collaboration with Professor Colin Aitken from Edinburgh University's Statistics department, along with the Home Office and Derbyshire Police. Colin was interested in applying statistical methods to forensic science, and subsequently founded the journal *Law, Probability and Risk*, serving as its inaugural chief editor.

The goal was to build statistical profiles of offenders using both data and the detectives' domain knowledge. It was a fascinating mix of hard data and human intuition. The results were published in a Home Office report titled "Predicting an Offender's Characteristics using Statistical Modelling", and Colin and I co-authored a paper on it as well [1].

ZL: I remember struggling with recovering the causal structure in my own PhD work. Pearl suggested the Chow-Liu algorithm, but our experiments with medical data were only partly successful. Still, it was a rich area to explore. What are your memories of those years in Edinburgh?

AG: Oh, I have very fond memories. There were several other projects - EPSRC, EU-funded - and the intellectual atmosphere was lively. But it wasn't just the research. The social life was rich too. I made many friends and colleagues in Scotland, and some of them have contributed to this Festschrift, which is lovely.

After ten years at Heriot-Watt, I moved on to Royal Holloway in 1993. Then, in 1996, I gave my Inaugural Lecture there. It was a chance to reflect on everything I'd done since arriving in the UK. That lecture is actually included as a chapter in this book.

4 Reliable Machine Learning, Royal Holloway University of London, 1993 - present.

ZL: You were appointed to the established chair in Computer Science at Royal Holloway in 1993 and also served as the Head of Computer Science department from 1995 to 2005. In 1998, the Centre for Reliable Machine Learning was established. What was that period like?

AG: It was a big transition. In Autumn of 1992, I was offered a job at Royal Holloway, University of London as a Professor of Computer Science at the University of London[5]. I started the new job on the 1st of September 1993 - it took a year to move the family from Edinburgh to London. The UK was in

[5] At the time, all professorial appointments were at the University of London (UL) rather than for individual Colleges of UL.

recession, and selling our house in Edinburgh wasn't easy. So the move was slow and a bit stressful.

Once I settled in, I dove into teaching and managing research grants. I also began organising a series of seminars on Bayesian inference with Unicom Seminars. Among the speakers were Vladimir Vapnik from AT&T, Judea Pearl from UCLA, Phil Dawid from UCL, David Spiegelhalter from Cambridge, and many others. Those seminars ran from 1993 to 1995 and eventually led to three edited volumes published by Wiley and Springer [5,7,4].

ZL: That's when you met Vapnik?

AG Yes, during those seminars. I invited him to join our department, and in 1995, he accepted a part-time professorship while continuing his work in the USA. Around that time, he published "The Nature of Statistical Learning Theory" and later "Statistical Learning Theory" [25,26]. These were landmark texts. They introduced Support Vector Machines—SVMs—which built on the Generalised Portrait method he'd developed with Alexey Chervonenkis[6], originally presented in their joint monograph [31].

The key innovation was the use of kernel functions within Hilbert-Schmidt theory. That allowed for powerful non-linear classification. Our department was one of the first to implement SVMs. We built a demo system for binary classification, and it really took off - downloaded by over 300 institutions worldwide.

4.1 Prediction with confidence; conformal predictors

ZL: Apparently, your and Volodya Vovk's interests in SVM and its limitations led to your work in the development of Conformal Predictors. Could you say a little bit about this link? And where did you first meet Volodya Vovk?

AG: I met Volodya at the EuroCOLT conference in Barcelona in 1995. Talking to him, you can immediately see that this unassuming and quietly spoken man is ingenious. He always finds an original point of view and converts his ideas into a rigorous and powerful theory. His mathematical instinct and brilliant technique led him to reconsider the foundations of classical probability theory, proposed by his University supervisor Andrey Kolmogorov, and to establish a new calculus and interpretation of probability [27,20].

Another area of research in which Volodya has been very active is machine learning, where he has also been pushing the boundaries in developing the Conformal Predictors and related methods. He also laid the game-theoretic foundations of the Prediction with Expert Advice method. I won't go into all his achievements here - there's not enough space - but I'll say this: he's a close friend, and working with him over the past 30 years has been a privilege.

AG: The idea of developing confidence measures really took shape as we explored both the strengths and the limitations of Support Vector Machines.

[6] A historical account of the Generalised Portrait can be found in [28].

Fig. 8. L-R: Vladimir Vovk, Alexander Gammerman, Vladimir Vapnik, 1996

That's what led us to Conformal Predictors - or CP for short. It's a statistically rigorous framework for quantifying predictive uncertainty, and it offers something quite rare in machine learning: provable validity.

Back in 1995-96, Volodya Vovk and I had a series of conversations with Vladimir Vapnik. We were discussing SVMs and Vapnik's broader concept of *transduction*- the idea of making predictions directly for new data points, rather than building a general model first. It was a subtle but powerful shift in thinking.

Then further discussions between Volodya and me happened while I was preparing my 1996 Inaugural Lecture [6] - it also includes a section on transduction. Subsequently, a paper was published [12] reflecting these discussions.

From there, the ideas evolved. Over the years, we formalised the theory and published it across a number of papers, talks, and books [29,19,10,30]. Conformal Prediction represents, in many ways, a paradigm shift in how we think about uncertainty in machine learning. It doesn't just give you a point prediction - it tells you how confident you can be in that prediction, with statistical guarantees.

That's something most machine learning methods still struggle to offer. And it's one of the reasons I believe CP has such lasting value.

ZL: For the benefit of our readers, could you summarise the basic features of Conformal Predictors?

AG: First of all, CP provides provably valid measures of confidence. It also does not make any additional assumptions about the data beyond the i.i.d. assumption - or even a weaker assumption of exchangeability.

What is also important that any known machine learning algorithm can be used to develop a CP. It will produce the well-calibrated prediction regions.

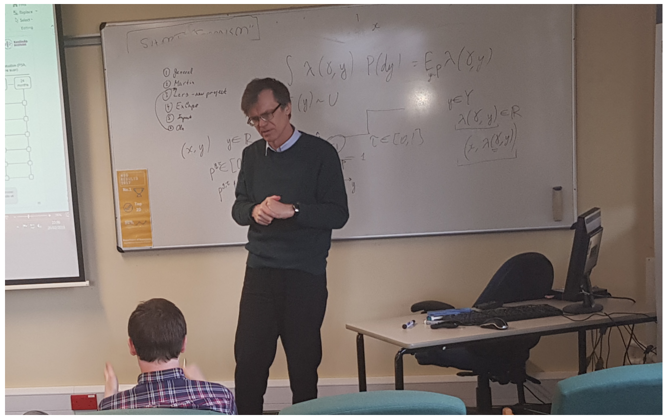

Fig. 9. Volodya Vovk, 2018

Several other techniques, we developed, that keep the validity property; they include *Venn-Abers reliable probabilistic* prediction; *Conformal Predictive Distribution*; *Conformal Testing Martingales* for testing the data to satisfy the i.i.d. assumption (and for detecting changes in distribution) and others.

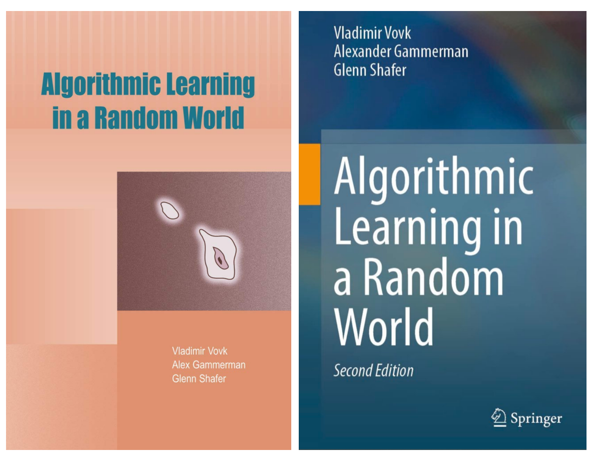

Fig. 10. Algorithmic Learning in a Random World: 2005 (first edition) and 2022 (second edition)
.

ZL: What role do think CP could play in the new and exiting research in Large Language Models?

AG: Recent development of *Deep Learning* and *Large Language Models* (LLMs) has already a considerable impact on the conformal predictors. CPs have evolved to address the needs of these powerful but often lacking calibration techniques [15]. I think that by enhancing calibration, interpretability, and trust, Conformal Predictors will play an increasingly critical role in making machine learning more reliable and statistically valid.

One persistent limitation of LLMs is their lack of a grounded notion of truth—they generate text, but they are not inherently reliable for factual accuracy and require strict verification. We often face this problem. We hope, however, that Conformal Prediction, with its rigorous treatment of uncertainty, might help instill in these models something resembling a statistical conscience.

4.2 Conformal Predictors Applications

ZL: Conformal Predictors have been successfully applied in many fields. Could you mention some of the most interesting ones from your point of view?

AG: Yes, indeed, we have been involved in many applications and grants from the UK, EU, China, Cyprus, etc. One of the latest projects was under the Amazon Research Awards programme entitled "Conformal Martingales for Change-Point Detection". There are applications in: Pharmaceutical industry (drug discovery); Medicine (diagnostics and treatment, including ovarian cancer, heart problems, gastroenterology, depression, etc.); and many others. Some of them can be found on our website: `https://cml.rhul.ac.uk/projects.html`.

For example, in a collaborative project with AstraZeneca (AZ), Conformal Predictors were employed to help prioritise which chemical compounds should be tested - to speed up the search for promising candidates in drug development. According to AZ report, this approach allowed to make a reduction of approximately one-third in both the time and cost associated with the discovery process [11,23].

While the pharmaceutical companies are understandably discreet about the financial implications of replacing traditional assays with predictive models. However, we estimated that such efficiencies could translate into savings ranging from $250 million to $1 billion per drug.

ZL: The results for AstraZeneca are impressive. Indeed, it is already clear that Conformal Predictors have a wide circle of applications and can have a significant societal impact.

5 Teaching and Supervision

ZL: Let's talk briefly about teaching and supervision. What was your experience like of teaching in both countries?

AG: When I first arrived in Britain, I hadn't done much teaching. In the USSR, teaching and research were kept separate—universities taught, research institutes did the research. The British model, where both happen under one roof, struck me as very sensible.

It's healthier, really. You're not constantly haunted by the pressure to publish. And from a teaching perspective, it's great—students get exposed to current research, not just dusty textbook material.

ZL: What courses did you teach? Did you use any particular teaching techniques?

AG: I taught across the board - from first-year undergraduates to Master's students. Courses ranged from theoretical topics like computational theory to hands-on programming in C++. There's always debate about what kind of mathematics computer science students should learn - predicate logic, analysis, discrete maths. Personally, I think the specific content matters less than the mindset. What's important is teaching students how to frame problems precisely, solve them logically, and interpret results clearly. That's the real goal.

ZL: Many students say you're great at explaining complex ideas - and that you often tell jokes or stories. Is that something you learned, or does it come naturally?

AG: I don't know. All I am trying to do is to pass on my enthusiasm for the subject. Teaching demands a skill set beyond research—it's a performance art. One must learn to engage, to play the role of lecturer-as-actor before a student audience. Obviously, a joke or anecdote can work: it relaxes the room, and makes learning a little bit easier.

ZL: You also taught in different countries while on your sabbaticals or as a visiting professor. Any interesting comparison with the UK teaching?

AG: Yes, I gave lectures - or full courses - in the USA, Spain, France and China between 1995 and 2011. In some places, students were very strong in mathematics but less confident in programming. That contrast made me think: perhaps we should be teaching more maths to our own students. I even wrote a report to the Royal Society about it, who sponsored some of my travels. But I'm not sure anything changed as a result.

ZL: And your main contribution to teaching?

AG: Definitely supervision. I've supervised many PhD students - sometimes solo, sometimes with colleagues - and it's been one of the most rewarding parts of my academic life. Watching students grow into independent researchers is a joy.

6 Administration

ZL: We've covered your research and teaching. Let's turn to your time as Head of Department at Royal Holloway. What was that experience like?

AG: In 1995, the Principal of Royal Holloway, Professor Norman Gowar, invited me to take on the role of Head of Department. At the time, Professor Chris Mitchell, who headed the department, had completed his five-year term and was ready to pass the baton. I was less than eager - my plate was already full - but Professor Gowar had a remarkable talent for persuasion. After several conversations, I agreed, on the condition that the College support the expansion of the Department and the creation of new research groups: one in Machine Learning, the other in Bioinformatics.

At that time, the department's reputation was largely built on its work in Information Security. Important work, of course, but I felt we needed to broaden our research base.

Fig. 11. With the staff members of Computer Science Department, Royal Holloway, 1998

ZL: And did the College support that vision?

AG: They did. With their backing, we gradually expanded. We built up a strong Machine Learning group and began assembling a promising Bioinformatics team. We recruited some excellent people - lecturers and professors with real energy and vision. Sadly, after 2005, the Bioinformatics direction was put on hold. Priorities shifted, and the momentum we'd built in Bioinformatics was lost. That was disappointing, but such is the nature of institutional life - winds change.

ZL: Beyond the day-to-day admin, you also launched some initiatives to raise the department's profile - like the Computer Learning Research Centre (renamed later to the Centre for Reliable Machine Learning - CRML) and the Kolmogorov Lecture and Medal.

AG: What the Department truly needed was greater visibility - both nationally and internationally. Our first major target was the 1996 Research Assessment Exercise (RAE), a periodic audit in which every university department in the UK is graded on its research output, with scores ranging from 1 to the coveted 5*.

I'd only just taken over, so we had very little time to prepare. But we worked hard—really hard—and managed to secure a grade 4. That was a solid result, given the circumstances.

Then came the 2001 RAE. This time, we were ready. We earned a grade 5, and the EducationGuardian ranked us 9th in the country out of more than 100 Computer Science departments. That was a proud moment for all of us.

ZL: Yes, I found the RAE 2001 Panel Feedback Report. It was very complementary about our ML research. Here is what the Report said: *"The Panel were impressed by the strength within the Computer Learning Group and the profound impact that their work has had on theory and applications"*.

AG: It was good to get appreciation for our work in ML.

Looking back, I do think our efforts helped raise the Department's profile—no question about that. But I've always had mixed feelings about the RAE, and later the REF. For all the time, energy, and strategic manoeuvring they demand, I remain unconvinced that they've genuinely improved the quality of research.

What they've certainly done is create a framework - a kind of bureaucratic ritual - for allocating funding and ranking departments. It gives administrators and funding bodies a sense of control, a veneer of objectivity. But whether it fosters real intellectual progress? I'm not so sure.

In some ways, it's like judging a symphony by the number of notes played. You can measure output, citations, impact factors - but the deeper qualities of research, the originality, the risk-taking, the long gestation of ideas - that's much harder to quantify. And often, it's precisely those things that get squeezed out when everyone's chasing metrics.

Still, we played the game. And we played it well. But I've always felt that the real work - the kind that matters - happens in the quiet corners, in conversations, in long walks with a notebook, and in the minds of people who aren't thinking about scores.

ZL: In 1998, the College established a new Machine Learning Centre. You were clearly instrumental in this - how did it come about, and what were your arguments?

AG: By 1997, it had become clear that our machine learning activities needed a more coherent structure. Research was branching into several promising directions - Statistical Learning Theory, Prediction with Expert Advice, Reinforcement Learning, Conformal Prediction, and others. The momentum was there, but without a dedicated hub, we risked fragmentation. Establishing a centre to consolidate and focus these efforts seemed both timely and necessary.

So I began conversations within the department and with the College leadership. We needed a dedicated hub - a centre that could bring everything together, give us focus, and signal to the wider world that we were serious about machine learning.

We drafted a proposal for the RHUL Academic Board and in February 1998, the Board approved the creation of the Computer Learning Research Centre (CLRC), which was later renamed the Centre for Reliable Machine Learning[7]. The Centre was staffed by a team of full-time academics (Volodya Vovk, Chris Watkins, John Shawe-Taylor and myself); there were some part-time academics, including Vladimir Vapnik (also at AT&T, USA), Alexey Chervonenkis (also at Institute of Control Sciences, Russia), and Glenn Shafer (also at Rutgers University, USA) and the visiting professors Chris Wallace (Monash University, Australia), Ray Solomonoff (Oxbridge Research, USA), Jorma Rissanen (IBM Almaden Research Centre, San Jose, USA), Leonid Levin (College of Arts and Sciences, Boston University, USA) and others.

If I'm not mistaken, back in 1998, there were only a handful of machine learning centres worldwide. Ours at RHUL, and Carnegie Mellon's, were among the very few.

ZL: The Centre has trained a number of PhD students over the years. How did it start?

AG: Over the years, the College supported us with a number of PhD studentships. This allowed us which allowed to train an excellent set of researchers. Among them were Ilia Nouretdinov, Jason Weston, Craig Saunders, and others. They completed their doctorates and have since gone on to lead machine learning research in major tech companies and universities. Today, you'll find them at Google, Meta, Microsoft, Amazon, and many other institutions around the world. It's been gratifying to see their work shape the field.

ZL: In the same year, 1998, you organised a colloquium with several world-leading academics and researchers. Could you describe the event and the invited speakers?

AG: Yes, in 1998 we marked the 30th anniversary of the Computer Science Department - founded in 1968 - with a series of colloquia. One of them was devoted to Machine Learning and Inference with the title *The Importance of*

[7] More information about the Centre for Reliable Machine Learning is available at `https://cml.rhul.ac.uk/`

Being Learnable. The event attracted nearly 100 participants from across the UK, Europe, the USA, Israel, and Australia, from academia and industry.

There were five invited speakers: Ray Solomonoff, Vladimir Vapnik, Alexei Chervonenkis, Jorma Rissanen, and Chris Wallace - towering figures in the fields of inductive inference and machine learning. Their work, dating back to the 1960s, helped ignite the intellectual revolution that shaped the modern theory of learning.

Fig. 12. Colloquium 1998: The Importance of Being Learnable. L-R: J. Rissanen, V. Vapnik, A. Gammerman, A. Chervonenkis, C. Wallace, R. Solomonoff

The originators of the VC theory, Vladimir Vapnik and Alexei Chervonenkis spoke on developments in statistical learning theory over the past 30 years. Jorma Rissanen outlined the theory of the Minimum Description Length principle and stochastic complexity. Chris Wallace, raised the fascinating open question of the relationship between different inductive approaches. Ray Solomonoff, discussed the application of many effective ML techniques to a wide variety of problem areas.

ZL: You were one of the founding organisers of the Kolmogorov Lecture and Medal. How did this initiative begin?

AG: It started in 2003, when we realised that Kolmogorov's centenary was approaching. It felt like the right moment to honour his legacy - not just as a towering figure in probability theory, but as someone whose ideas continue to shape modern mathematics and computer science.

So we launched the University of London Kolmogorov Lecture and Medal. The very first recipient was Ray Solomonoff, one of the pioneers of artificial intelligence[8] and co-creator of Kolmogorov-Solomonoff complexity [21]. It was a fitting tribute.

Over the years, the event grew into something quite special. We welcomed an extraordinary lineup of speakers - Per Martin-Löf, Leonid Levin, Jorma Rissanen, Yakov Sinai, Robert Merton, Vladimir Vapnik ... each one brought their own perspective, and each had, in some way, built on the intellectual foundations that Kolmogorov laid[9].

We certainly enjoyed those meetings. The conversations were rich, sometimes surprising, and always deeply rooted in the kind of thinking Kolmogorov inspired—rigorous, curious, and beautifully abstract.

ZL: You were also involved in organising a conference called COPA - Conformal Prediction with Applications. How did it come about?

AG: In 2011, during a conversation with our former PhD student Harris Papadopoulos, I floated the idea of launching a conference dedicated to Conformal Predictors. The concept quickly gained interest, and by 2012 we had organised the inaugural symposium, Conformal and Probabilistic Prediction with Applications (COPA), held in Greece, as part of the AIAI conference. And so, COPA was born.

Since then, it has grown into an independent annual conference in the machine learning calendar. This year marks the 14th COPA, which will be held in London. Accepted papers have been published in the Lecture Notes in Artificial Intelligence, and since 2017, in the Proceedings of Machine Learning Research (PMLR), reflecting the symposium's continued academic impact and evolution.

ZL: Thank you again for taking the time to talk with me and share your experiences and insights, Alex. There is a summary of your work on your personal webpage `www.gammerman.com` and Wikipedia `https://en.wikipedia.org/wiki/Alexander_Gammerman`.

It has been a wonderful opportunity to learn about your fascinating research journey from the corridors of Soviet research institutes to British academia, spanning from 20th to 21st centuries! From physics to mathematical biology, to pattern recognition, and ultimately to conformal prediction - your journey has been fascinating. Please carry on!

AG: Thank you! Indeed, research is a wonderful thing. But as Albert Einstein said once: *If we knew what it is that we are doing, it wouldn't be called research.* [Laughs].

[8] Ray Solomonoff was among the pioneering minds at the 1956 Dartmouth Conference where the term "Artificial Intelligence" was first coined. He joined RHUL as a Visiting Professor in 1997.

[9] Further details are available at: `http://kolmogorov.cml.rhul.ac.uk/`.

Fig. 13. Kolmogorov Lecture 2005. L-R: A. Chervonenkis, V. Vovk, R. Solomonoff, Per Martin-Lof, A. Gammerman, Z. Luo

References

1. C.G.G. Aitken, A. Gammerman, G. Zhang, T. Connolly, D.B. Bailey, R. Gordon, and R. Oldfield. Bayesian belief networks with an application in specific case analysis. In *Computational Learning and Probabilistic Reasoning*, pages 169–184. John Wiley & Sons, 1996.
2. H. Borthwick. The high-energy light action controlling plant responses and development. *Proc Natl Acad Sci U S A*, 64(2):479–486, 1969.
3. A. Gammerman. Correlation between processes in phytochrome and photomorphogenic reactions of plants. *Doklady Biophysics, Proceedings of the Academy of Sciences of the USSR*, 226:13–15, 1976.
4. A. Gammerman. *Probabilistic Reasoning and Bayesian Belief Networks.* Alfred Waller, Henley-on-Thames, 1995.
5. A. Gammerman. *Computational Learning and Probabilistic Reasoning.* John Wiley & Sons, Chichester, 1996.

6. A. Gammerman. *Machine Learning: Progress and Prospects.* Inaugural Lecture Series, Royal Holloway University of London, 1996.
7. A. Gammerman. *Causal Models and Intelligent Data Management.* Springer-Verlag, 1999.
8. A. Gammerman and L. Fukshansky. A mathematical model of phytochrome - the receptor of photomorphogenic processes in plants. *The Soviet Journal of Developmental Biology*, 5:122–129, 1974.
9. A. Gammerman and R. Thatcher. Bayesian diagnostic probabilities without assuming independence of symptoms. *Methods Inf Med*, 30(1):15–22, 1991.
10. A. Gammerman and V. Vovk. Hedging prediction in machine learning (with discussion). *The Computer Journal*, 50:151–170, 2007.
11. A. Gammerman, V. Vovk, and Z. Luo. Novel machine learning methods for improving drug discovery efficiency, 2021. Royal Holloway University of London REF 2021 Impact Report.
12. A. Gammerman, V. Vovk, and V. Vapnik. Learning by transduction. *UAI'98: Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, 14:148–155, 1998.
13. A.Y. Gammerman and L. Fukshansky. Theoretical analysis of processes in the phytochrome pigment system. *Studia Biophysica*, 52(1):65–72, 1975.
14. A.Y. Gammerman and L. Fukshansky. Theory and calculations of dynamics of phytochrome transformations in the green leaf. *Soviet Plant Physiology*, 18, 1976.
15. P. Giovannotti and A. Gammerman. Calibrated large language models for binary question answering. *Proceedings of Machine Learning Research*, 230:218–235, 2024.
16. M.I. Lanin, A. Gammerman, and I.P. Sheetova. Development of a set of programs using information-theoretic methods for solving systematic problems. *Abstracts of reports of the 12th International Botanical Congress*, 1:27, 1975.
17. M.I. Lanin, A. Gammerman, I.P. Sheetova, and O.D. Bykov. Application of multidimensional information-theoretic analysis in the study of the contingency of qualitative features. volume 61, pages 99–104. VIR, 1978.
18. J. Pearl. *Probabilistic Reasoning in Intelligent Systems.* Morgan-Kaufmann, 1988.
19. C. Saunders, V. Vovk, and A. Gammerman. Transduction with confidence and credibility. *Proceedings of the 16th International Joint Conference on Artificial Intelligence, Stockholm, Sweden, 1999*, 1999.
20. G. Shafer and V. Vovk. *Game-theoretic foundations for probability and finance.* Wiley, Hoboken, NJ, 2019.
21. R. Solomonoff. Universal coding, information, prediction, and estimation. *IEEE Transactions on Information Theory*, 30(4):629–636, 1984.
22. A.R. Thatcher. Relationships between Bayesian and confidence limits for predictions. *Journal of the Royal Statistical Society: Series B*, 26(2):176–192, 1964.
23. P. Toccaceli and A. Gammerman. Combination of inductive mondrian conformal predictors. *Machine Learning*, 108:489–510, 2019.
24. A. Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London Series B*, 237(641):37–72, 1952.
25. V. Vapnik. *Nature of Statistical Learning Theory.* Springer, 1995.
26. V. Vapnik. *Statistical Learning Theory.* Wiley, 1998.
27. V. Vovk. A logic of probability with application to the foundations of statistics (with discussion). *Journal of the Royal Statistical Society B*, 55:317–351, 1993.
28. V. Vovk, A. Gammerman, and H. Papadoupolus. *Measures of Complexity. Festchrift in honor of Alexey Chervonenkis.* Springer, 2015.

29. V. Vovk, A. Gammerman, and C. Saunders. Machine learning applications of algorithmic randomness. *Sixteenth International Conference on Machine Learning*, pages 444–453, 1999.
30. V. Vovk, A. Gammerman, and G. Shafer. *Algorithmic learning in a random world; 1st edition 2005; 2nd edition 2022.* Springer.
31. В. Вапник and А. Червоненкис. *Теория распозновaния образов.* Москва, 1974.
32. А.П. Малинин, С.А. Сафонова, and Ю.А. Пунанов. *Развитие детской онкологии в Ленинграде и Санкт-Петербурге.* Санкт-Петербург 2015.
33. М.Г. Пименов, Р.М. Малкина, and А.Я. Гаммерман. Анализ Внутривидовой Химической и Морфологической Изменчивости (adenostyles rhombiofolia (willd.) m.pimen. Методами Многомерного Информационного Анализа и Автомотической Группировки. *Растительные Ресурсы*, 17(1):23–36, 1981.