SANSKRIT AND AI / PART 06 / STUDY NOTES

Sanskrit and AI, Part 6: Rules Before Training

Study notes by Sanskrit AI Universe, based on this video

Video by Dr P.R.K. Varma, Founding Editor · DrVARMA PRK · 17 Nov 2026

Not the creator's own slides

About these notes

— Official title: Sanskrit and AI part 6 : Rules Before Training | Dr.P.R.K.Varma.,M.D.,D.M.

— Speaker and channel: Dr P.R.K. Varma, Founding Editor · DrVARMA PRK

— Published 17 Nov 2026 · Running time 29 min 44 s · https://www.youtube.com/watch?v=ONCr7kzBpk0

— These pages are study notes drawn from the video's own title, description and captions, plus the journal's already published Watch note. They are not a transcript and not a substitute for watching the film.

— The journal treats videos as leads, not as sources you should cite on their own. Check the papers and editions behind a claim.

— Caption source: English auto-generated subtitles. Wording may be imperfect. Distinctive phrases follow the captions closely.

This part in brief

— On the road, he saw that the answer could not be right, and he turned back.

— Because this part is about what must be settled before learning begins and about what happens when it is not.

— Everyone wants to do the model work, not the data work.

— Then comes 75 sūtras that do not form a single word.

— They give names, and they say how every later rule is to be read.

Chapter one, the inventory

— Three questions about what goes in, which labels exist, how clean the material is, and how many times the same thing is counted.

Problem one picture, one label

— More than one image in five showed objects from several classes, and the official label was often not what people saw as the main subject.

— The scheme was somebody had decided before any training that each thing belongs to exactly one class.

— A class earns a name only when some rule needs it.

— Scholars have worked out why some rules need her to stand with the voice sounds, others need it with the hissing sounds at the far end.

Problem two, errors at the source

— The answer key itself was faulty, and the authors showed that with only a little more noise, corrected labels would put a smaller model ahead of a larger one. larger one.

— Meter is the feet, ritual the hands, astronomy, the eye, etymology, the ear, grammar is the mouth, and Shiksha, the science of pronunciation is the nose.

— A manual of pronunciation that carries Pāṇini's name then describes the good reciter.

— The standard for the raw material was fixed and taught before the first rule of grammar.

Problem three

- More than one line in a hundred of what it wrote, unprompted, was copied straight from its training text.
- More than four test items in every 100 overlapped with the training text.
- To speak of two men named Rama, Sanskrit does not say Rama and Rama.
- So the rule removes the repetition and keeps the number, one form with its count attached. count attached.

Chapter two, the conventions

— Four questions about reading and instruction, how far a label reaches, what happens where the instruction is silent, what to do when there is only one, and which goes with which.

Problem four, how far does a label reach

- A model trained on their work learns four rules at once and nobody chose any of them. of them.
- So the suffix comes after the root G to conquer and gives Jaya, victory.
- When a rule names a unit, whatever has been inserted inside that unit is covered too.
- When a rule names a word stem, its feminine form is covered as well.

Problem five, where the instruction is silent

— They trained the same image model 50 times, the same design, the same data, the same score on the test.

— Many of his rules replace one thing with another, and many of them do not say which part is replaced.

— One rule tells the pronoun that to take the sound It does not say where.

— By default, only the final sound gives way and becomes Second, if a rule says after this, the change falls on the first sound of what follows.

Problem six when there is only one

— In 2023, researchers took a standard test for code written by language models.

— Pass rates fell by as much as 19 to 29% and the ranking of the models changed.

— Pāṇini says what happens at the edge where there is only one, treat it as the first and as the last.

— The rule applies and the form comes out right The same line covers the other end where a rule speaks of the first sound, a lone sound counts as the first.

Problem seven, which goes with which

— A translation system learns from pairs, a sentence in one language and the same sentence in another.

— Sentences set beside the wrong partner, text in the wrong language, collections filed under the wrong language name.

— The lists were long and nobody had checked that the first item of one answered the first item of the other.

— Sūtra 1 3 10 3 10 First with first, second with second, stated once and relied on everywhere.

Chapter three, the regiment

— Three questions about the learner, in what order to teach, how to examine, and when the training is over.

Problem eight, the order of the lessons

- But when training time was short or the data were noisy, easy first clearly helped. clearly helped.
- So order matters most when the learner is under pressure.
- The order in which knowledge is stored need not be the order in which it is taught.
- The Aṣṭādhyāyī is arranged for the rules, not for the reader.

Problem nine, a test you can prepare for

— In 2024, a team wrote more than a thousand new arithmetic problems by hand in the style of a famous public test.

— Wherever the needle stopped, that page was opened and the candidate had to explain it. explain it.

— The needle could stop anywhere, so the only preparation was to know the whole book. whole book.

— Our accounts of these tests are late and come through tradition rather than records, but the design is clear and it is sound. is sound.

Problem 10, the first good answer

— Build the next round of tests from that failure and decide before training begins who has the right to say enough.

— Put the 10 answers together as a list to go through before training begins.

— That is what Pāṇini's opening pages are, decisions written down before the work begins.

— The phoneticians were training human voices and the Upaniṣad is about the self, not about machines. about machines.

Sources and credit

— Study notes by Sanskrit AI Universe, based on the video “Sanskrit and AI part 6 : Rules Before Training | Dr.P.R.K.Varma.,M.D.,D.M.”. These are not the creator's own slides.

— Video by Dr P.R.K. Varma, Founding Editor, posted on the YouTube channel DrVARMA PRK.

— Watch the video: https://www.youtube.com/watch?v=ONCr7kzBpk0

— Captions used: English (auto-generated) (en), retrieved on 24 November 2026 via a public Invidious instance. Auto-captions can mis-hear words; obvious proper-name spellings were given in IAST where the journal already uses them.

— Also used: the official video title and description, and the Watch catalogue note already published in this journal (presenter and why the video is listed).

— If the Founding Editor later supplies his own deck for this part, replace the PPTX and PDF in src_assets/slides/ (see the journal README, “How to swap in RK's own slides”).