

1/10/24

Team name: Rangers

Team members: Layla Barela, Jack A., Alexandra Sosa

Sponsor: Ashli Knoell

Using AI to keep an endangered language alive

Problem Definition

Did you know today only 170,000 people can speak the Navajo language, and only a quarter of them can speak proficiently. The Navajo language and many other northern indigenous languages are endangered and dying. 65 indigenous languages are already extinct and 75 are nearing extinction. Many of native americans stopped speaking their indigenous languages due to the 1900's when native american kids were sent to boarding schools to be rid of traditional American Indian ways of life and have them replaced with mainstream American culture. These children who were brought to these boarding schools often didn't teach their children their indigenous language resulting in fewer and fewer speakers of their language.

The Navajo language is only on three language learning apps and they only cover the very basics of the language. Many language learning apps fail to go in depth in the language due to its complexity. Did you know in the Navajo language there are 48 handling verbs that change if the item is at rest, an opened object, closed object, or an object filled with other objects? You can be saying two completely different words just depending if you use a high or low tone. Described by linguists Robert W. Young and William Morgan as a "hopeless maze of irregularities," its unique grammar, syntax, and tonal pronunciation are so indecipherable to outsiders that, during the Second World War, the U.S. Marines used Navajo as a form of military code. Most of the time the only way to learn conversational Navajo is to live on a reservation where the language is spoken.

The Plan

Our team plans to create a Siri like AI that speaks Navajo. With our project we hope to make an AI that helps you learn how to ask questions and conversate in the Navajo language. Our team plans to use machine learning and create a language model for the Navajo language. We will use the NVIDIA Nemo framework to build our AI model. We plan to divide our project into different parts. We plan to collect and train the data on one person. Once we have our model made we want the AI to perform an action according to the word spoken. After we've succeeded in having the AI perform the action for us we want the AI to be able to translate what was spoken in Navajo to English and have the words appear as English text on screen so that the speaker can understand what was spoken.

Current Progress

We experimented a little with Deepspeech and are currently collecting data. Our original plan was to use Deepspeech as our framework but we decided to use NVIDIA Nemo once we realized Deepspeech wasn't being worked on anymore. Since this is a new field for us our group is working to understand how machine learning works.

Expected results

Because of the complexity of the Navajo language, it's impractical to expect our AI model to be able to completely understand the Navajo language, However we believe that our model will be able to understand and translate basic Navajo commands, in turn teaching the user of the model. While you can't learn a whole new language just from telling an AI what to do, we hope that our project can help accelerate and fill the gaps in learning conversational Navajo.

Works Cited

Brinkhof, Tim. "Why Navajo is one of the most difficult languages on Earth." *Big Think*, 27 November 2023, <https://bigthink.com/high-culture/navajo-language/>. Accessed 10 January 2024.

"Chapter 3: Boarding Schools - Native Words, Native Warriors - National Museum of the American Indian." *National Museum of the American Indian*, <https://americanindian.si.edu/nk360/code-talkers/boarding-schools/>. Accessed 10 January 2024.

Goodman, Sylvia. "Native American Languages Are Disappearing. Colleges Could Help Preserve Them." *The Chronicle of Higher Education*, 23 January 2023, <https://www.chronicle.com/article/native-american-languages-are-disappearing-colleges-could-help-preserve-them>. Accessed 10 January 2024.

Miller, Jeremy. "Most native tongues of the West are all but lost." *High Country News*, 2 March 2015, <https://www.hcn.org/issues/47.4/most-native-tongues-of-the-west-are-all-but-lost>. Accessed 10 January 2024.

Woody, Jalen, et al. "Data shows huge reduction in Diné speakers." *Navajo Times*, <https://navajotimes.com/reznews/data-shows-huge-reduction-in-dine-speakers/>. Accessed 10 January 2024.