

“Ask and you shall receive”: Duke students use machine learning to optimize use of business intelligence

Durham NC—April 23, 2022—An interdisciplinary group of six Duke students launched “SEBI”, a project that aims to revolutionize the way managers access and use data to lead their decision-making process. Combining their backgrounds on Computer Science, Statistical Statistics and Economics, the students developed a tool that allow managers to orally request a piece of data or information they require and receive it immediately, without having to file requests to the data analysis team or manually access the database. The SEBI, Speech Enabled Business Intelligence, uses an artificial intelligence model that recognizes spoken language and converts it to text, then another model identifies what is requested in the plain text, converts it into a proper query, and prints out the requested information. This innovative tool represents a practical implementation of the phrase “ask and you shall receive”.

As the project developers remarked, “in the competitive world we live in today, companies’ success requires the use of data to guide them to optimal decisions. However, we identified that managers that do use data as an input usually relied on rigid reports or dashboards with a very limited set of indicators to guide their decisions.” Via SEBI, mid and high-level managers are granted immediate access to updated business information at a question’s length away. Furthermore, the implications of the project are not restricted to officials from quantitative departments nor limited to business and market indicators. By eliminating the requirement of using statistical or data querying software, SEBI knocks the entry barrier to data use, enabling heads of not-so data-driven departments, like Human Resources or Planning, to inform their decisions by simply asking questions like “what are the top complaints of personnel with less than 5 years in the company?”.

“Implementing SEBI can also improve business performance through an indirect channel: the optimization of resources from the data analytics department”, added Hank, one of the team leaders. Now that managers have an expedited path of access to data, the number of daily requests to be processed and answered by the data analytics team reduces drastically, freeing key human resources to devote time to more relevant departmental projects. “Through these two mechanisms – the optimization of access to data and the re-direction of efforts of the data analytics team- SEBI enables the transformation to an innovative, data-driven, and, thus, more profitable enterprise”, he concluded.

The members of the project combined their multiple backgrounds with pilot tests to construct a useful tool for a wide range of enterprises and settings. “Being a regional manager, I am constantly traveling to oversee processes and results. So, I rarely have time to sit down and look for the data. With this tool, I can use my smartphone’s voice assistant and get the required information instantly.” said one of the participants in the pilot phase. Other participants

highlighted as a key element that they were able to tailor the set of indicators, statistics, and output format of the results to their preference. SEBI is the first product from this promising group of students and is already generating important improvements in the economy, contributing to their goal of constructing a more efficient and data-driven world.

FAQs

Q1. How does minimum viable product work?

A1. SEBI is an AI (Artificial Intelligence) driven speech assistant for business intelligence which helps you find the right information quickly and easily. The final product is a voice enabled service which can be integrated on existing cloud infrastructure being used by a company. The SEBI service performs two independent functions which are critical to automated business intelligence: (i) Convert user speech to English language text using speech recognition (ii) Convert text to SQL query using language translational models. We plan to extend the current product to support speech enabled output via two additional enhancements: (i) Convert results from resulting SQL query to text using 'table to text' generation models (ii) Convert resulting text to speech using TTS (Text to speech) models. The SEBI service interacts with users through voice assistants and lets users ask business relevant questions which can be answered by gathering and summarizing data across various data sources within an organization. The above mentioned four critical functions convert user speech to appropriate SQL query and subsequently convert SQL results to speech to communicate business insights with the user. SEBI enabled voice assistants are installed in the meeting rooms within the office premises of our clients and can be operated by legitimate users with appropriate data access controls.

See technical details for model details and preliminary results

Q2. Which cloud services can work with SEBI?

A2. SEBI can be integrated with a company's usual preferred cloud partner. The product currently has a seamless integration with four cloud providers: Google Cloud and Amazon Web Service (AWS), Microsoft Azure and Apple iCloud. This allows SEBI to access companywide data sources stored on cloud across various business segments and geographies, thus extending a unified experience for the users. We also provide customized solutions for integrating SEBI with alternative cloud providers such as Samsung Cloud.

Q3. Do I need to get a physical voice assistant like Alexa to start using SEBI service?

A3. SEBI requires a physical voice assistant like Alexa, Google Assistant, Microsoft Cortana, or Apple Siri to provide voice enabled business intelligence. JayKay Technologies is currently dedicated to enhancing user experience for SEBI and relies on specialized hardware developed by Amazon, Google, Microsoft, and Apple for voice assistant.

Q4. How does SEBI charge for the service?

A4. We operate on a subscription-based pricing model, allowing customers to use as many devices as possible with SEBI and pay on a regular basis for the SEBI service. Additionally, SEBI uses cloud resources which would entail additional computing expenses charged independently by the cloud service provider. The detailed description of our subscription plans is available on our website.

Q5. How does SEBI address data privacy concerns for information sharing across geographies and business segments?

A5. SEBI relies on cloud providers' fine-grained access control and their compliance with local data protection laws to engage users in a data secured environment. SEBI uses transformer-based attention models to establish user identity through speech recognition. Once the user identity is established, the access to sensitive information on cloud is controlled by access controls such as Identity and Access Management (IAM) service on AWS. This allows users access sensitive information across business verticals depending on their seniority level in the organization. Additionally, cloud providers' compliance with local data protection laws ensures legally compliant information and data sharing between geographies.

Q6. How does SEBI ensure data security over voice enabled communication?

A6. SEBI establishes user identity through speech recognition for secured business intelligence. Companies are required to regularly update voice samples for legitimate users along with their updated business roles for secured access and better data access controls. Additionally, the transformer-based attention models for establishing user identity are trained using GAN (Generative Adversarial Network) which ensures adversarial training defense against identity theft attack. This eliminates the possibility of unintended access to an unauthorized person or computer-generated voice, pretending to be a legitimate user.

Q7. What business and system assumptions are being made while developing this product?

A7. One of our biggest assumptions is that users are always interested in getting a summary of structured tabular data. Though the product in existing form deals with structured datasets across variety of sources, it is not empowered to deal with any intelligible task on unstructured data. Currently, we do not anticipate much demand for business intelligence on unstructured data but miscalculating on this assumption might impact future user growth and product penetration.

Q8. How big is the potential market for SEBI and the opportunity cost of not pursuing this product?

A8. SEBI is a B2B product, and we expect it to be commercially viable for companies which have at least 3000 corporate employees. We expect a significant uptake among companies which have at least 5000 employees. Assuming 4% penetration in the prospective market segment of employers with at least 5000 employees, we estimate potential market for SEBI in the US at \$3 billion.

Q9. Are there any legal ramifications for misinterpreted questions by SEBI?

A9. JayKay Technologies has currently rolled out a beta version for SEBI which requires customers to sign a waiver absolving it of liability in the event of misinterpreted questions by SEBI. However, we expect to have legal backing for the product from the next commercial release.

Q10. How is the speech to text system trained

A10. The speech to text model we used is a pre-trained wav2vec2 model from Facebook. The base model is pretrained and fine-tuned on 960 hours of Librispeech on 16kHz sampled speech audio.

Q11. How is the speech to text system evaluated

A11. Word Error Rate (WER) is commonly used to evaluate the performance of a speech recognition system and we apply WER to evaluate our model. The model achieves 3.4% WER on the clean dataset and 8.6% on the other dataset.

Q12. How is the text to SQL system trained

A12. We use language translational models to convert questions in English language to SQL queries. We used Wiki SQL dataset of size 360k made available by Salesforce, to train the model.

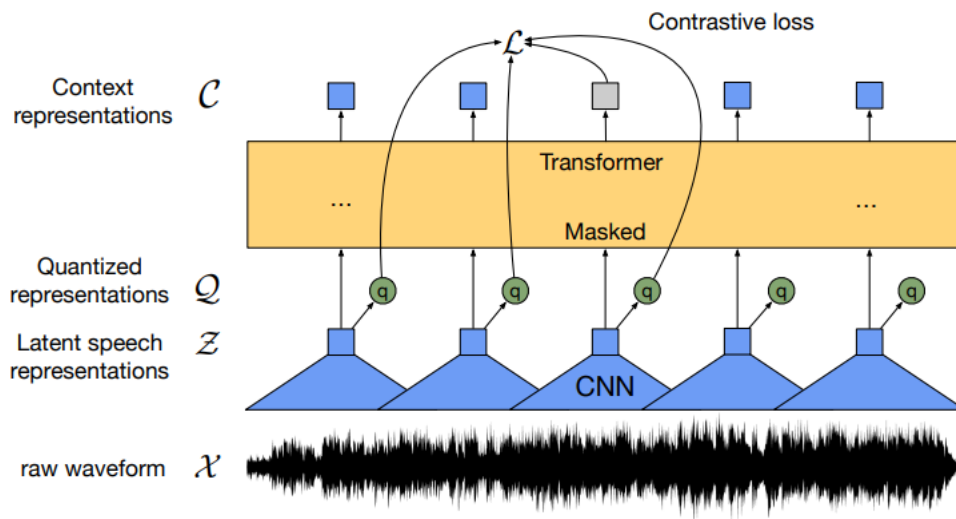
Q13. How is the text to SQL system evaluated

A13. The results from SQL query are compared with truth labeled English questions from Wiki SQL dataset and evaluated for accuracy. We targeted accuracy of atleast 80% as evaluation criteria while training the model.

Technical Specifications

Speech to Text Model: simulation results

A high-level overview of the model architecture for speech to text conversion is shown below. It uses state of the art transformer based models to convert speech to text with high accuracy.



The table below contains an implementation of the speech to text model implemented. As commented in the FAQ section, the model takes an audio file (first column), processes it, and gives a text as output (third column). This output can be compared to the second column with the expected translation as shown below.

Audio File	English translation of original English Speech	Output from the model
0000.wav	HOW MANY HEADS OF THE DEPARTMENT ARE OLDER THAN FIFTY-SIX	HOW MANY HEARS OF THE DEPARTMENT ARE OLDER THAN FIFTEEN SIX
0001.wav	LIST THE NAME, BIRTHDAY AND AGE OF THE HEADS OF DEPARTMENTS ORDERED BY AGE	LACE THE NAME BOARESTATE AND AGE OF THE HAS OF DEPARTMENTS ORDERED BY AGE

Text to SQL: simulation results

The output from Speech to tech model is English Questions which are translated to SQL queries using Text to SQL model. Below we present sample English questions and their translation into SQL queries by the model.

English Questions	SQL outputs from model
What is terrence ross' nationality	SELECT col2 FROM table WHERE col0 = Terrence ross
What club was in Toronto 1995-96	SELECT col5 FROM table WHERE col4 = 1995-96

We present below the input tabular data on which SQL operation is performed after converting English question to SQL query.

col0	col1	col2	col3	col4	col5
Player	No.	Nationality	Position	Years in Toronto	School/Club Team
aleksandar radojević	25	serbia	center	1999-2000	barton cc (ks)
shawn respert	31	united states	guard	1997-98	michigan state
quentin richardson	n/a	united states	forward	2013-present	depaul

Appendix

Git Link for details: [1999Lyd/test-to-sql-model-implementation \(github.com\)](https://github.com/1999Lyd/test-to-sql-model-implementation)

Academic research used for model implementation and calibration:

1. wav2vec 2.0: A Framework for Self-Supervised Learning of Speech Representations
2. Hybrid Ranking Network for Text-to-SQL