

अहो YouTube, या व्हडिओमध्ये मी आपल्याला दर्शवणार आहे की आपण व्हॉसिपर नावाच्या प्लायथनमधील वनिमूल्य मुक्त स्त्रोत पॅकेजचा वापर करून कोणत्याही ऑडिओला मजकूरात द्रुतपणे कसे रूपांतरित करू शकता.मी हे स्थापित केले आहे हे दर्शवणार आहे, मी ते कसे चालविले याचे एक उदाहरण दर्शवा आणि विद्यमान लायब्ररीशी तुलना करा.म्हणून प्रारंभ करून, आपण कदाचित येथे पहात असलेल्या व्हॉसिपर गेट हब रेपॉजिटरीवर जायचे आहे आणि आपण ते कसे स्थापित करू शकता याबद्दल ते सूचना देतात.जेव्हा आपण फक्त नावाची कुजबुज स्थापित करता तेव्हा आपण योग्य आवृत्ती स्थापित करणार नाही हे लक्षात ठेवण्याची आता एक गोष्ट.आम्हाला या गिट रेपॉजिटरीमधून स्थापित करायचे आहे.म्हणून फक्त ही पपि इंस्टॉल कमांड घ्या आणि आपल्या वातावरणात चालवा की आपण प्लायथन चालवित आहात.आणि त्यांनी येथे नमूद केले आहे की आपल्याला एफएफएम पीईजी स्थापित करणे आवश्यक आहे.हे करण्यासाठी काही सूचना आहेत, परंतु माझ्या संगणकावर ते आधीपासूनच स्थापित केले आहे.आता मी व्हॉसिपर स्थापित केले आहे, तर आपण हे चाचणी करू शकेन असे काही ऑडिओ बनवूया.म्हणून मी काही मुहावरे म्हणणार आहे.मॉडेलसला समजण्यासाठी सहसा मुहावरे कठीण असतात.जरी हे फक्त मजकूरासाठी भाषण आहे.ही एक प्रकारची मजेदार असेल.मला क्लाऊड 9 वर एक युक्ती म्हणून पोनी म्हणून रहायला आवडेल जे माशीला दुखापत होणार नाही.मी पाण्याबाहेर मासे असल्यासारखे आणि हिवाळ्याच्या खाली असलेल्या फडिलितकेच फटि आहे.चला हे जतन करूया.चला लाट म्हणून जतन करूया.एकदा ते स्थापित झाल्यानंतर आम्ही कमांड लाइनमधून हे कसे चालवू शकतो याबद्दल त्यांच्याकडे सूचना आहेत.ते येथे दाखवणारे प्लायथन एपीआय कसे वापरावे हे मी तुम्हाला दर्शवित आहे.तर हे खरोखर सोपे आहे.आम्ही फक्त कुजबुज आयात करतो.मग आम्ही आमचे मॉडेल तयार करणार आहोत, जे आम्ही लोड करणार आहोत.मॉडेल ज्याला बेस म्हणतात.आणि नंतर फक्त हा मॉडेल ऑब्जेक्ट वापरून, आम्ही आमच्या ऑडिओ फाईलवर लिप्यंतरण चालवितो.म्हणून मी त्याचे नाव दिले.चला वेव्ह आवृत्ती वापरू.आम्हाला हा नकाल परत करायचा आहे.आता मी हे लक्षात घेतले की जेव्हा मी हे धावले तेव्हा कुडाच्या अर्ध्या टेन्सर आणि फ्लोट टेन्सरमुळे मला ही त्रुटी मिळाली.मी हे सोडवण्यात सक्षम होतो.तर हे लक्षात ठेवण्यासारखे काही तरी आहे.जर ते आपल्यासाठी कार्य करत नसेल तर आपल्याला कदाचित फ्लोटिंग पॉईंट 16 पडण्यसाठी सेट करण्याची आवश्यकता असू शकेल.आणि आपण येथे धावल्यानंतर आपण पाहू शकता, त्यास इंग्रजी म्हणून आधीपासूनच भाषा आढळली आणि नंतर या परिणामी ऑब्जेक्टमध्ये त्यामध्ये काही भिन्न पद्धती आहेत, परंतु आम्हाला या आतून काय मिळवायचे आहे ते फक्त एक मजकूर आहे आणि आम्ही ते पाहू शकतो की ते दिसते आहेजसे परिणाम चांगला आहे त्याप्रमाणे मला क्लाऊड नऊवर एक युक्ती म्हणून पोनी म्हणून रहायला आवडेल ज्यामुळे माशीला दुखापत होणार नाही मी पाण्यातून मासे बनवितो आणि यामुळे पाण्यातून या मासे थोड्या वेळाने गडबड झाली फडिल म्हणून तंदुरुस्त आहे आणि कदाचित मी हे स्पष्टपणे सांगितले नाही की हे जाणून घेण्यासाठी आणि एक गोष्ट स्पष्टपणे म्हणाली नाही की जेव्हा आपण प्रथम हे चालवित तेव्हा ते बेस मॉडेल डाउनलोड करावे लागेल.तर कदाचित आपणास एक प्रगती बार जाताना दिसेल आणि आपल्याला ते मॉडेल डाउनलोड करावे लागेल.आणि असे म्हणते की जेव्हा आपण हे लिप्यंतरण चालवित तेव्हा ते आपल्या ऑडिओ फाईलच्या 30 सेकंदांचा भाग घेत आहे आणि त्यावर अंदाज चालू आहे.आता आपण घेऊ शकता असा आणि एक दृष्टीकोन देखील आहे, जो नमिन् स्तरीय दृष्टीकोन आहे, जथी आपण प्रत्यक्षात मॉडेल तयार करता आणि नंतर आपण ऑडिओ ऑब्जेक्ट आणि निमुना ट्रिम तयार करता.तर आपण फक्त हे सुनिश्चित करा की हा ऑडिओ भाग फक्त 30 सेकंद आहे.सेकंद क्वी ते 30 सेकंदांसह थाप देईल कारण मॉडेलची इनपुट म्हणून अपेक्षित असलेल्या लांबीची लांबी आहे.मग ते लॉग माउस स्पेक्ट्रोग्राम बनवित आहे.हे भाषा शोधत आहे आणि आम्ही येथे डीकोड करू शकतो आणि इच्छित असल्यास बरेच अधिक प्रयास प्रदान करू शकतो.जर मी हा सेल चालविला तर पुन्हा ही त्रुटी मिळवा, जी मी आता डिकोडिंग प्रयासामध्ये सेट करू शकतो, एफपी 16 फॉल्ट्स समान आहे.आणि प्रत्यक्षात

यावेळी असे दसिते की हे सर्व काही ठीक झाले आहे.मी पाण्यातून मासे सारखे असावे.आणिएक फा डल म्हणून तंदुरुस्त आहे.तर ते कुजबुजण्यासाठी आहे.मला फक्त त्याची तुलना वदियमान प्रकारच् या मॉडेलशी करायची आहे.आणअसे करण्यासाठी एक लोकप्रयि लायब्ररी म्हणजे स्पीच रकिग्नशि शन लायब्ररी.आम्ही स्पीच रकिग्नशि न लायब्ररी चालवण्याचा मार्ग म्हणजे आम्ही ते आयात कर तो आणनंतर ही ओळखकर्ता ऑब्जेक्ट तयार करतो, जी आम्ही नंतर आमची ऑडिओ फाईल लोड करू शकतो.त्यानंतर, आपण ओळखकर्ता ऑब्जेक्ट घेऊ शकता आणतियासाठी काही भन्नि ओळ खण्याच्या पद्धती आहेत.आणआम्ही Google ओळखू वापरणार आहोत आणपिरणिम काय आ हे ते पाहूया.तर असे दसिते की त्याने कोणतेही वरिमचनिहे जोडले नाहीत आणढिग नऊ भन्नि आहेत .मला फ्लायला इजा होणार नाही अशा एका युक्तीच्या पोनी म्हणून क्लाऊड नऊ वर रहायला आवडे ल.परंतु एक गोष्ट लक्षात ठेवण्याची एक गोष्ट म्हणजे ती प्रत्यक्षात Google स्पीच रकिग्नशि न एपीआय वापरत आहे.व्हस्पर लायब्ररी, आपल्याकडे प्रत्यक्षात मॉडेल डाउनलोड केले आहे आ णतें वापरण्यासाठी आपले आहे.या कोडसह प्रसदिध झालेल्या व्हस्पर पेपरवर एक नजर टाकण्य ाची मी शफारस करतो.मॉडेल कसे प्रशक्षिती केले गेले आणतें वापरलेल्या आर्कटिक्चरबद्दल ते तपशीलवार देखील जातात.व्हस्पर वेगवेगळ्या भाषांच्या गुच्छावर कार्य करते.ते म्हणतात त्या का मगरींमध्ये भाषेच्या आधारे बदलते.तर आपण येथे गीथब रेपोवर जाऊ शकता जथें त्यांच्याकडे एक प्लॉट आहे की येथे बारसाठी कोणत्या भाषा प्रत्यक्षात उत्कृष्ट कामगरी करतात हे दर्शवति आहे .लहान चांगले आणभोठे म्हणजे ते अधिक वाईट करते.म्हणून हे मॉडेल ज्या भाषांवर कार्य करते त् या भाषांची संख्या अजूनही प्रभावी आहे.