**Project Overview: Vaccine Hesitancy, Social Media Influence, and Misinformation in Alaska**

**1. Introduction and Hypothesis**

Vaccine hesitancy remains a significant public health issue, particularly in rural and underserved communities. Social media plays a pivotal role in shaping public perceptions, spreading both factual information and misinformation. This study hypothesizes that vaccine hesitancy in Alaska is significantly influenced by social media discourse, misinformation trends, and socio-cultural factors, which vary across demographic groups.

Additionally, the study evaluates the effectiveness of AI-driven RAG models in summarizing public sentiment and detecting misinformation. By comparing findings from social media sentiment analysis, interview sentiment analysis, and AI-generated responses, we aim to identify key drivers of hesitancy and develop strategies to improve vaccine confidence.

**2. Datasets and Methodology**

This research employs a mixed-methods approach, integrating quantitative sentiment analysis, qualitative interview insights, and AI-driven text generation models.

**2.1 Datasets**

The study utilizes two primary datasets:

1. **Social Media Dataset (Twitter)**
   * **Source:** Kaggle dataset filtered to include tweets originating from Alaska.
   * **Size:** ~1300 tweets after pre-processing.
   * **Processing Steps:**
     + **Geolocation Extraction:** Filtered tweets to include only those from **Alaska** with city-level location details.
     + **Sentiment Classification:** Applied **RoBERTa-based** NLP models to categorize sentiments.
     + **RAG:** Identified some key insights by querying 41 carefully selected questions.
2. **Interview Dataset**
   * **Participants:** 87 individuals from targeted households and stakeholder groups.
   * **Data Type:** Transcribed responses to structured vaccine-related questions.
   * **Focus Areas:**
     + Vaccine safety concerns.
     + Trust in healthcare institutions.
     + Influence of family and community narratives.

**2.2 The 41 Core RAG Questions**

The study's 41 questions focus on understanding vaccine hesitancy, misinformation, and public sentiment, grouped into six themes.

**Theme 1: Vaccine Hesitancy Reasons**

1. What are the most common reasons for COVID-19 vaccine hesitancy among Alaskans?
2. What are the key concerns about COVID-19 vaccine safety and effectiveness in Alaska?
3. How do past negative experiences with vaccines shape vaccine hesitancy?
4. What role do family and peer networks play in influencing vaccine hesitancy?
5. How do rural and urban Alaskans differ in their reasons for vaccine hesitancy?

🔹 **Goal from this theme:** These questions help uncover why Alaskans hesitate to vaccinate, how concerns are influenced by past experiences, and how different communities (rural vs. urban) perceive vaccines.

**Theme 2: Misinformation & Conspiracy Theories**

1. What are the most common misinformation themes about COVID-19 vaccines in Alaska?
2. How do Alaskans respond to fact-checking efforts on vaccine misinformation?
3. How do religious beliefs shape vaccine hesitancy in Alaska?
4. What conspiracy theories about COVID-19 vaccines are being discussed in Alaska?
5. How has vaccine misinformation changed over time in Alaska?
6. How do social media and traditional news sources differ in spreading vaccine misinformation?
7. What misinformation about vaccine side effects is most commonly discussed in Alaska?

🔹 **Goal from this theme:** These questions help identify misinformation trends, track the evolution of false narratives, and evaluate how fact-checking efforts impact public trust.

**Theme 3: Social & Cultural Factors**

1. What cultural factors influence vaccine confidence among Native Alaskan communities?
2. How do vaccine hesitancy attitudes vary between different socio-economic groups in Alaska?
3. How do Alaskan parents view childhood vaccinations compared to COVID-19 vaccines?
4. How do rural Alaskans perceive vaccines differently than urban residents?

🔹 **Goal from this theme:** These questions explore how culture, socio-economic status, and geographic location influence vaccine trust and hesitancy.

**Theme 4: Barriers to Vaccination (Beyond Social Media)**

1. What are the biggest logistical challenges Alaskans face in getting vaccinated?
2. How do vaccine access issues in rural Alaska compare to urban centers like Anchorage?
3. What concerns do Alaskans express about the speed of vaccine development and approval?
4. How do people in Alaska perceive government vaccine mandates or incentives?
5. What role does transportation, cost, and availability play in vaccine hesitancy?
6. How has the availability of COVID-19 vaccines changed vaccination behavior in Alaska?

🔹 **Goal from this theme:** Understanding practical barriers (e.g., cost, travel distance, supply issues) can inform policy recommendations to increase vaccine uptake.

**Theme 5: Social Media Solutions & Policy Recommendations**

1. What strategies have been most successful in engaging vaccine-hesitant individuals in Alaska?
2. What role should social media platforms play in countering vaccine misinformation in Alaska?
3. What social media campaigns have been most effective in increasing vaccine confidence in Alaska?
4. What policies or interventions have been proposed to improve vaccine uptake in Alaska?
5. What lessons can be learned from successful vaccine campaigns in Alaska?
6. What incentives or outreach methods have been effective in overcoming vaccine hesitancy in Alaska?

🔹 **Goal from this theme:** These questions identify effective strategies for public health messaging and social media interventions.

**Theme 6: Healthcare Professionals & Stakeholder Insights**

1. What do healthcare professionals in Alaska believe are the biggest drivers of vaccine hesitancy?
2. How do healthcare professionals address vaccine concerns in patient interactions?
3. What are the biggest challenges that healthcare professionals face in increasing vaccine acceptance?
4. How do healthcare professionals in Alaska perceive the impact of social media on vaccine decisions?
5. What strategies do public health officials recommend to improve vaccine trust in Alaska?

🔹 **Goal from this theme:** These questions offer expert perspectives on vaccine hesitancy and how public health professionals can improve vaccination rates.

**2.3 Methods**

The study applies three analytical approaches: sentiment analysis, misinformation tracking, and AI-based Q&A retrieval.

**2.3.1 Sentiment Analysis**

**Objective:** Categorize vaccine-related opinions from tweets and interviews into positive, negative, or neutral sentiments.

**Pipeline:**

* **Model:** **CardiffNLP/twitter-roberta-base-sentiment**
* **Processing Steps:**
  + Tokenized tweets using Hugging Face's AutoTokenizer.
  + Sentences were truncated to 512 tokens for consistency.
  + Classified sentiment using a fine-tuned RoBERTa model.

**Urban vs. Rural Analysis:**

* **Location classification** was performed based on tweet geolocation, categorizing tweets into urban (e.g., Anchorage) vs. rural (e.g., Nome, Bethel, Kotzebue) sentiment distributions.

**2.3.2 Retrieval-Augmented Generation (RAG) Model Analysis**

**Objective:** Use AI models to generate answers to **41 core vaccine-related questions** using retrieved social media discussions.

**A. Llama-Index Pipeline**

* **Model:** **Llama-2-7B-chat-hf (Meta)**
* **Embeddings:** **sentence-transformers/all-mpnet-base-v2**
* **Indexing Method:** Vector Store Index using **FAISS**.
* **Retrieval Method:**
  + Tweets were embedded using **SentenceTransformer**.
  + **Similarity-based retrieval** was performed to find top 100 matching tweets per query.
* **Answer Generation:** Used Llama-2 to synthesize responses from retrieved text.

**B. T5-based RAG Pipeline**

* **Model:** **T5-base** (Hugging Face Transformer)
* **Retrieval Method:** FAISS-based vector retrieval (top-5 similar tweets per query).
* **Answer Generation:** T5 model was prompted to generate summary responses.

**Comparison of RAG Models:**

* **Llama-2-7B** provided **contextually rich and nuanced answers**.
* **T5-based model** was **less detailed** and sometimes generated **repetitive outputs**.

**3. Key Findings**

**3.1 Sentiment Analysis (Social Media)**

* **Negative Sentiment Dominance:** A large portion of tweets expressed **concerns over vaccine safety, mandates, and personal freedoms**.
* **Recurring Misinformation Themes:**
  + Claims that COVID-19 vaccines alter DNA.
  + Fear of long-term health risks despite scientific evidence.
  + Belief that natural immunity is better than vaccines.
* **Political and Regional Divide:**
  + Rural Alaskans exhibited more resistance to vaccines due to government distrust.
  + Urban residents had more balanced sentiments but expressed concerns about mandates.

**3.2 Interview Sentiment Analysis**

* **Top Vaccine Concerns:**
  + Fear of side effects and unknown risks.
  + Mistrust in government and pharmaceutical companies.
  + Influence of family and peer networks.
* **Urban vs. Rural Perceptions:**
  + Rural communities were hesitant due to autonomy concerns.
  + Urban respondents were more likely to trust public health officials.

**3.3 RAG Model Insights**

* The Llama-2 model generated detailed, context-aware responses.
* The T5 model was more generic and sometimes incoherent.
* Both models effectively identified misinformation trends, validating their potential for automated misinformation tracking.

**4. Discussion and Implications**

**4.1 Social Media’s Role in Vaccine Perceptions**

* **Echo Chambers Reinforce Hesitancy:** Algorithmic filtering intensifies misinformation circulation.
* **Community-Based Approaches Needed:** Health campaigns must engage trusted local figures.
* **Challenges in Fact-Checking:** Many users reject corrections as “government control.”

**4.2 AI’s Role in Public Health Research**

* **Llama-2 RAG models** show promise in detecting misinformation patterns in real-time.
* **Future Potential:** Deploying AI-based chatbots for vaccine education.

**5. Recommendations**

**5.1 Public Health Strategies**

* **Localized Messaging:** Rural vs. urban narratives require customized health messaging.
* **Community-Led Interventions:** Engage faith leaders, community elders, and local healthcare providers.

**5.2 AI & Social Media Interventions**

* **Real-time Misinformation Tracking:** Expand Llama-2 RAG models to detect false claims.
* **Social Media Collaboration:** Work with platforms to fact-check viral misinformation.

**6. Conclusion**

This study confirms that social media significantly influences vaccine hesitancy in Alaska, with misinformation, trust issues, and socio-cultural differences driving vaccine concerns. AI-driven RAG models effectively track misinformation trends, highlighting their potential for real-time public health interventions.

Future research should explore scaling AI-powered misinformation monitoring and longitudinal tracking of vaccine sentiment shifts.