QUESTION:

From your understanding, which model will work better for the sequence tagging problem? Logistic regression or CRFs? Why? You can give examples if you and them helpful for explanation!

ANSWER:

Logistic regression vs CRF:

Logistic Regression:

The Logistic regression classifier, when predicting the label yi, can use any features of the input sequence x and the position i. Here, we think about the features as being computed for the observed input. It does not consider the neighbouring input samples as CRF.
In other words, a logistic regression model, assumes complete independence between the different components yi in y.

CRF:

Conditional random fields are an extension of logistic regression that incorporates sequential information in the labels, while still supporting the use of arbitrary features. This is obviously more powerful than Logistic regression when it comes to sequence tagging problem.
The variabilities and dependencies of the token to other given tokens are also considered when CRF tagger is used.
Another way to put this is that a simple classifier predicts a label for a single sample without regard to neighbouring input samples. However, a CRF does take this into account.

Comparison for performance in Sequence tagging:

According to the above details and understanding, CRF should do a better job in sequence tagging problem. Since we are tagging a sequence of words, or tagging the words considering the context and the sequence, a CRF should always out perform a logistic regression classifier in this case.
What would logistic regression classifier do?
For the above, logistic regression would just see a word as a token and give features according to it, and use them to classify results.

However, a CRF will also see the dependencies of a specific token with others, and give the results accordingly. With this, it also does what a Log-reg classifier does, that is also taking into account simple features token wise.

Example:

Suppose the word to be tagged is get is “Manchester United”. Here Manchester would be classified a s a location. But, we know that Manchester United Is a Football team, and hence should be classified as a Sports team. Given the categories logistic regression classifier would tag Manchester as a location. However, CRF would look at the context, and the dependencies of the token, to give the result as sports team.

The CRF would look at united and it's features and give the results as a sports team, for "Manchester United." That is how it should work, by considering dependencies and by using Viterbi algorithm to find the best fit.

So CRF would see the features and outputs of the previously tagged inputs and then predict the answer using Viterbi algorithm.

Another unique example would be to determine the nationality of a festival. We are just presented with a simple festival detail additional information. Without context, it's hard to determine where it comes from because many festivals have similar kind of celebration and rituals. Now we are presented with a few more festival names from the same country. Let's say the additional festivals are the ones, we could easily figure out belong to which country. The pattern can now be recognized, and we realize that the festival is probably belonging to India (as looking at other festivals like Diwali, Holi). This is how CRF works. Instead of just blindly looking at something, it learns about the context before making any decision.